

Run on:	August 25, 2005, 03:18:03 ; Search time 166 Seconds (without alignments)	382.101 Million cell updates/sec	26	421	49.1	99	8	ADR22067	Anti - Hepa
Title:	US-10-664-038-1		27	421	49.1	99	8	ADR21937	Anti - Hepa
Perfect score:	857		28	420	49.0	99	8	ADR21997	Anti - Hepa
Sequence:	1 ARILNLKKKTNVTPTVAHRT.....ALPGPWRMASGFNKATAMQQ 164		29	419	48.9	99	8	ADR21978	Anti - Hepa
Scoring table:	BLOSUM62		30	417	48.7	99	8	ADR21944	Anti - Hepa
	dapop 10.0 , Gapext 0.5		31	417	48.7	139	6	ABU61880	HCV core
Searched:	2105692 seqs, 386760381 residues		32	416	48.5	99	8	ADR21959	Anti - Hepa
Total number of hits satisfying chosen parameters:			33	416	48.5	99	8	ADR21948	Anti - Hepa
Minimum DB seq length:	0		34	415	48.4	99	8	ADR22033	Anti - Hepa
Maximum DB seq length:	20000000000		35	415	48.4	99	8	ADR22072	Anti - Hepa
Post-processing:	Minimum Match 0% Maximum Match 100%		36	415	48.4	99	8	ADR22065	Anti - Hepa
Database :	A_Geneseq_16Dec04:*		37	415	48.4	99	8	ADR21947	Anti - Hepa
	1: geneseqD19808:*		38	415	48.4	99	8	ADR22029	Anti - Hepa
	2: geneseqD19908:*		39	415	48.4	99	8	ADR21993	Anti - Hepa
	3: geneseqD20008:*		40	415	48.4	99	8	ADR21950	Anti - Hepa
	4: geneseqD20018:*		41	414	48.3	99	8	ADR22070	Anti - Hepa
	5: geneseqD20028:*		42	414	48.3	99	8	ADR1945	Anti - Hepa
	6: geneseqD2003a8:*		43	414	48.3	99	8	ADR22038	Anti - Hepa
	7: geneseqD2003b8:*		44	414	48.3	99	8	ADR21900	Anti - Hepa
	8: geneseqD20048:*		45	413	48.2	99	8	ADR21986	Anti - Hepa
ALIGNMENTS									
RESULT 1									
			ID	ADG22790	standard; protein; 160 AA.				
			XX	ADG22790;					
			XX	AC					
			DT	26-FEB-2004	(first entry)				
			XX	DE					
			XX	Hepatitis C virus F protein.					
			XX	KW	virucide; antiviral RNA silencing pathway modulator; gene inactivation; RNA silencing; recombinant construct; antiviral RNA silencing pathway; viral infection; hepatitis C virus; F protein; RNA silencing suppressor.				
			XX	OS					
			XX	Heptatitis C virus.					
			XX	PN	US2003219407-A1.				
			XX	PD	27-NOV-2003.				
			XX	PP	15-MAY-2002; 2002US-00150283.				
			XX	PR	15-MAY-2002; 2002US-00150283.				
			XX	PA	(REGC ) UNIV CALIFORNIA.				
			XX	PI	Ding S, Li H, Li W;				
			XX	DR	WPI; 2004-021947/02.				
			XX	DR	N-PSBB; ADG22791.				
			XX	PT	New recombinant DNA construct comprising a polynucleotide sequence of a virus which activates RNA silencing, and a polynucleotide sequence of a gene, useful for inactivating gene in a cell.				
			XX	PT	The invention describes a recombinant DNA construct for inactivation of a gene in a cell. The construct comprises: a polynucleotide sequence of a virus which activates RNA silencing; and a polynucleotide sequence of the gene. The recombinant constructs are useful in inactivating a gene in a cell or in an animal. Modulators of antiviral RNA silencing pathway are useful for treating or preventing viral infection. This is the amino acid sequence 160 AA.				
			XX	PS	Disclosure; SEQ ID NO 1; 20pp; English.				
			XX	CC	The invention describes a recombinant DNA construct comprising a polynucleotide sequence of a gene in a cell. The construct comprises: a polynucleotide sequence of a virus which activates RNA silencing; and a polynucleotide sequence of the gene. The recombinant constructs are useful in inactivating a gene in a cell or in an animal. Modulators of antiviral RNA silencing pathway are useful for treating or preventing viral infection. This is the amino acid sequence 160 AA.				
			XX	CC	Query Match	97.7%	Score 837;	DB 8;	Length 160;
Result No.	Score	Query Match	Length	DB	ID	Description			
1	83.7	97.7	160	8	ADG22790	Adg22790 Hepatitis AAY44545 Hepatitis Abu61864 HCV core			
2	81.3	94.9	198	3	AAY44545	AAY44545 Hepatitis Abu61864 HCV core			
3	80.6	94.0	161	6	ABU61864	Abu61864 HCV core			
4	76.8	89.6	161	6	ABU61866	Abu61866 HCV core			
5	76.3	89.0	161	6	ABU61865	Abu61865 HCV core			
6	64.5	75.3	198	3	AAY44558	AAY44558 Hepatitis Abu61870 HCV core			
7	62.4	72.8	161	6	ABU61870	Abu61870 HCV core			
8	58.3	68.0	198	3	AAY44554	AAY44554 Hepatitis Abu61868 HCV core			
9	55.2	64.4	143	6	ABU61868	Abu61868 HCV core			
10	54.8	63.9	143	6	ABU61872	Abu61872 HCV core			
11	54.7	63.8	143	6	ABU61871	Abu61871 HCV core			
12	54.0	63.0	143	6	ABU61867	Abu61867 HCV core			
13	54.0	63.0	143	6	ABU61873	Abu61873 HCV core			
14	54.0	63.0	143	6	ABU61874	Abu61874 HCV core			
15	53.4	62.3	143	6	ABU61875	Abu61875 HCV core			
16	50.8	59.3	198	3	AAY44557	AAY44557 Hepatitis AAY44557 Hepatitis Abu61869 HCV core			
17	50.1	58.5	198	3	AAY44552	AAY44552 Hepatitis Abu61879 HCV core			
18	48.7	56.8	154	6	AAY44556	AAY44556 Hepatitis AAY44556 Hepatitis Abu61878 HCV core			
19	46.4	54.1	198	3	AAY44555	AAY44555 Hepatitis AAY44555 Hepatitis Abu61869 HCV core			
20	44.1	51.5	198	3	AAY44553	AAY44553 Hepatitis AAY44553 Hepatitis Abu61877 HCV core			
21	42.8	49.9	139	6	ABU61869	Abu61869 HCV core			
22	42.6	49.7	99	8	ADR22073	Adr22073 Anti-hepa			
23	42.5	49.6	99	8	ADR21979	Adr21979 Anti-hepa			
24	42.4	49.5	154	6	ABU61878	Abu61878 HCV core			
25	42.2	49.2	99	8	ADR21949	Adr21949 Anti-hepa			



be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain.

Sequence 161 AA;

Query Match 94.0%; Score 806; DB 6; Length 161;  
Best Local Similarity 99.4%; Pred. No. 7.56-74;  
Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
SQ 7 KCKKINVTPTVAHTSSRVAVSLSVEFTCRAGLDWTCARRRLPSGRNLEYDVSISPR 66  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 8 QKCTINVTPTVAHTSSRVAVSLSVEFTCRAGLDWTCARRRLPSGRNLEYDVSISPR 67  
Qy 67 LVGRAGAPLSPGTLGSPNAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIWRSIPLR 126  
Db 68 LVGRAGAPLSPGTLGSPNAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIWRSIPLR 127  
Qy 127 AASPTSWGTYRSSAPLLELPGPKWMAASCFWKTA 160  
Db 128 AASPTSWGTYRSSAPLLELPGPKWMAASCFWKTA 161

RESULT 4

ABU61866 standard protein; 161 AA.

XX ABU61866;  
AC XX

XX 23-OCT-2003 (revised)  
DT 14-AUG-2003 (First entry)

XX HCV core protein frameshift protein p17 #3.

XX HCV; hepatitis C infection; RNA frameshift; core protein; p17; viricide;  
KW hepatotropic; overlapping open reading frame; p21c; vaccine.  
XX Hepatitis C virus; genotype 1a.

OS XX US2002076415-A1.  
PN XX 20-JUN-2002.

XX PP 14-DEC-2000; 20000US-00736959.  
XX PR 14-DEC-1999; 99US-0170835P.

XX DR 2003-479366/45.

XX PS 99US-0170835P.

XX PA (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX XX DR WPI; 2003-479

CC acid encoding p17), an anti-viral composition (used to treat hepatitis C)  
 CC comprising a compound that binds to p17, antibodies directed against an  
 CC HCV core protein which are elicited by immunising an animal using the  
 CC partially purified protein p17, a method for analysing an HCV antigen in  
 CC a sample using the anti-p17 antibodies and detection of anti-HCV  
 CC antibodies in a sample using the p17 proteins. The Hcv p17 and the DNA  
 CC sequences that encode it may be used as vaccines for immunising patients  
 CC against HCV infection. The antibodies and the antiviral compound may also  
 CC be used for treating HCV infections. Hcv p17 and the antibodies may also  
 CC be used in immunoassays for detecting HCV antigens and/or antibodies in  
 CC samples for the diagnosis of HCV infections. The present sequence  
 CC represents a p17 protein from an HCV strain. (Updated on 23-Oct-2003 to  
 CC standardise OS field)

XX Sequence 161 AA;

Query Match 89.0%; Score 763; DB 6; Length 161;  
 Best Local Similarity 9.8%; Pred. No. 1.8e-69;  
 Matches 146; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Qy 7 KAKCNVTPVTAHRTSSSRVAVRSLEFTCRAGLDWVARRBLPSGNLEVDVSLSPR 66  
 Db 8 QRKENVTPVTAHRTSSSRVAVRSLEFTCRAGLDWVARRBLPSGNLEVDVSLSPR 67  
 Qy 67 LVGRPAGPGLSPGTGSPAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWRSSTPLR 126  
 Db 68 HVGRPAGPGLSPGTGSPAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWRSSTPLR 127  
 Qy 127 AASPTSGWGTYRSSAPLLEALPGPWRMASGFWKTA 160  
 Db 128 AASPTSGWGTYRSSAPLLEALPGPWRMASGFWKTA 161

RESULT 6

AAV44558  
 ID AAV44558 standard; protein; 198 AA.  
 XX  
 AC AAV44558;  
 XX  
 DT 04-APR-2000 (first entry)

Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;  
 KW prevention; diagnosis; therapeutic target; anti-HCV therapy;  
 KW HCV infection; DNA isolate D50482.  
 XX  
 OS Hepatitis C virus.

XX Key Location/Qualifiers

PT Misc-difference 139  
 PT /label= unknown  
 PT /note= "Corresponds to stop codon"  
 PT Misc-difference 143  
 PT /label= unknown  
 PT /note= "Corresponds to stop codon"  
 PT Misc-difference 161  
 PT /label= unknown  
 PT /note= "Corresponds to stop codon"  
 PT Misc-difference 178  
 PT /label= unknown  
 PT /note= "Corresponds to stop codon"  
 PT Misc-difference 184  
 PT /label= unknown  
 PT /note= "Corresponds to stop codon"  
 XX W09963941-A2.  
 XX 16-DEC-1999.  
 PD XX 09-JUN-1999; 99WO-US012929.  
 PP XX 09-JUN-1998; 98US-00886705.

CC PR 11-JUN-1998; 98US-00886705.

CC XX (BRAN/) BRANCH A D.

CC PA (WALE/) WALEWSKI J L.  
 CC (STUMP/) STUMP D D.

XX PI Branch AD, Walewski JL, Stump DD;  
 XX DR WPI; 2000-126431/11.

CC Novel Hepatitis C virus peptides useful in vaccine compositions, for  
 CC diagnosing HCV infection and as therapeutic agents.  
 XX PT Example 1; Page 39-41; 50PP; English.

XX The present sequence is a Hepatitis C virus (HCV) protein derived from an  
 CC HCV isolate AAB50482. The novel HCV protein is not encoded by the  
 CC standard HCV open reading frame but by an alternate reading frame (ARF)  
 CC which is +1 or +2 to the standard HCV ORF. The position of the first  
 CC nucleotide of the ARF may vary slightly depending upon the isolate. The  
 CC protein elicits an immune response in patients infected with HCV and are  
 CC produced during HCV infection. The present sequence is used in vaccine  
 CC compositions for preventing HCV infection. It is also used for diagnosing  
 CC HCV infection and as a target for anti-HCV therapy

XX Sequence 198 AA;

Query Match 75.3%; Score 645.5; DB 3; Length 198;  
 Best Local Similarity 76.4%; Pred. No. 2.2e-57;  
 Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARILNKKTKTNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60  
 Db 1 AQLNLCKPKNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60

Qy 61 VSLSPRLVGPRAGPGLSPGTGLPSMAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWVR 120  
 Db 61 DNLSPRPAGPAGPGLSPGTGLPSMAMRVNGQDGSCHPGALGLVGAAPRTPGVGRIVWVR 120

Qy 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 164  
 Db 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 165

XX Sequence 198 AA;

Query Match 75.3%; Score 645.5; DB 3; Length 198;

Best Local Similarity 76.4%; Pred. No. 2.2e-57;  
 Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARILNKKTKTNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60  
 Db 1 AQLNLCKPKNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60

Qy 61 VSLSPRLVGPRAGPGLSPGTGLPSMAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWVR 120  
 Db 61 DNLSPRPAGPAGPGLSPGTGLPSMAMRVNGQDGSCHPGALGLVGAAPRTPGVGRIVWVR 120

Qy 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 164  
 Db 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 165

XX Sequence 198 AA;

Query Match 75.3%; Score 645.5; DB 3; Length 198;

Best Local Similarity 76.4%; Pred. No. 2.2e-57;  
 Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARILNKKTKTNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60  
 Db 1 AQLNLCKPKNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60

Qy 61 VSLSPRLVGPRAGPGLSPGTGLPSMAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWVR 120  
 Db 61 DNLSPRPAGPAGPGLSPGTGLPSMAMRVNGQDGSCHPGALGLVGAAPRTPGVGRIVWVR 120

Qy 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 164  
 Db 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 165

XX Sequence 198 AA;

Query Match 75.3%; Score 645.5; DB 3; Length 198;

Best Local Similarity 76.4%; Pred. No. 2.2e-57;  
 Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARILNKKTKTNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60  
 Db 1 AQLNLCKPKNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60

Qy 61 VSLSPRLVGPRAGPGLSPGTGLPSMAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWVR 120  
 Db 61 DNLSPRPAGPAGPGLSPGTGLPSMAMRVNGQDGSCHPGALGLVGAAPRTPGVGRIVWVR 120

Qy 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 164  
 Db 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 165

XX Sequence 198 AA;

Query Match 75.3%; Score 645.5; DB 3; Length 198;

Best Local Similarity 76.4%; Pred. No. 2.2e-57;  
 Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARILNKKTKTNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60  
 Db 1 AQLNLCKPKNTPTVTAHRTSSSRVAVRSLEFTCRAGLDWVARRRLPSGRNLVEVD 60

Qy 61 VSLSPRLVGPRAGPGLSPGTGLPSMAMRAAGRDGSCLPVALGLGAPOTPGVGRAIWVR 120  
 Db 61 DNLSPRPAGPAGPGLSPGTGLPSMAMRVNGQDGSCHPGALGLVGAAPRTPGVGRIVWVR 120

Qy 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 164  
 Db 121 SSIPPLRASPTSGWGTSSAPLLEAPGPWMASGFWKTA-TMQQ 165

DR WPI; 2003-479366/45.  
 XX Isolated hepatitis C virus (HCV) proteins formed by expression of of  
 PT overlapping open reading frames in the core protein gene sequence through  
 PT a frame shifting mechanism, useful for vaccinating against, and detecting  
 PT HCV infections.  
 XX  
 PS Claim 9; Page 15; 37pp; English.  
 XX  
 CC The invention relates to an isolated and purified protein of the  
 CC hepatitis C virus (HCV) that is formed by expression of an overlapping  
 CC open reading frame in the core protein gene sequence through an RNA frame  
 CC shifting mechanism. The protein is termed p17 (the full length, unshifted  
 CC protein being p21c). Also included are a vaccine (including a DNA  
 CC vaccine) for immunising a mammal against hepatitis C (producing a  
 CC protective antibody) comprising at least 1 protein of p17 (or a nucleic  
 CC acid encoding p17), an anti-viral composition (used to treat hepatitis C)  
 CC comprising a compound that binds to p17, antibodies directed against an  
 CC HCV core protein which are elicited by immunising an animal using the  
 CC partially purified protein p17, a method for analysing an HCV antigen in  
 CC a sample using the anti-p17 antibodies and detection of anti-HCV  
 CC antibodies in a sample using the p17 proteins. The HCV p17 and the DNA  
 CC sequences that encode it may be used as vaccines for immunising patients  
 CC against HCV infection. The antibodies and the antiviral compound may also  
 CC be used for treating HCV infections. HCV p17 and the antibodies may also  
 CC be used in immunoassays for detecting HCV antigens and/or antibodies in  
 CC samples for the diagnosis of HCV infections. The present sequence  
 CC represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 161 AA;  
 Query Match 72.8%; Score 624; DB 6; Length 161;  
 Best Local Similarity 77.3%; Pred. No. 2.6e-55;  
 Matches 119; Conservative 7; Nismatches 28; Indels 0; Gaps 0;  
 XX  
 Qy 7 KKCCTNVTPTVAHRTSSRVAVRSVLVEFTCRAAGALDWYCARRERLPSGRNLVEYDVSLSPR 66  
 Db 8 QRKPNVTPTVAHRTSSRVAVRSVLVEFTCRAAGAPGWVACRLGTLPSGRNLVEGDNLSPR 67  
 XX  
 Qy 67 LVGDRAGPGLSPGTLGPSMAMRAAGGRDGSCLPYAHLGAGAQPQPGVGRAIWYRSSIPIR 126  
 Db 68 LADPRAAGPGLSPGTLGPSMAMRALGQDGSCHPAAPGLVGAAPTPGVGRIVIWRSSIPSH 127  
 XX  
 Qy 127 AASPTSWGTTTSSAPPLLEALPGPWRMASGEFWKTA 160  
 Db 128 AASPTSWGTTTSSAPPLLEALPGPWRMASGEFWKTA 161  
 XX

RESULT 8  
 AAY44554 standard; protein; 198 AA.  
 XX  
 AC AAY44554;  
 XX  
 DT 04-APR-2000; (first entry)  
 XX Hepatitis C virus protein encoded by DNA isolate D14853.  
 XX  
 KW Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;  
 KW prevention; diagnosis; therapeutic target; anti-HCV therapy;  
 KW HCV infection; DNA isolate D14853.  
 XX  
 OS Hepatitis C virus  
 XX  
 ID AAY44554 standard; protein; 198 AA.  
 XX  
 AC AAY44554;  
 XX  
 DT 04-APR-2000; (first entry)  
 XX Hepatitis C virus protein encoded by DNA isolate D14853.  
 XX  
 KW Hepatitis C virus protein; RNA frameshift; core protein; p17; virucide;  
 KW hepatotropic; overlapping open reading frame; p21c; vaccine.  
 XX  
 OS Hepatitis C virus; genotype 1b.  
 XX  
 FH Location/Qualifiers  
 FT Misc-difference 143  
 /label= "Corresponds to stop codon"  
 FT Misc-difference 161  
 /label= unknown  
 FT Misc-difference 184  
 /label= "Corresponds to stop codon"  
 FT

PT PT /label= unknown  
 PT PT /note= "Corresponds to stop codon"  
 XX XX  
 PN WO9633941-A2.  
 XX XX  
 PD 16-DBC-1999.  
 XX XX  
 PP 09-JUN-1999; 99WO-US012929.  
 XX PR 09-JUN-1998; 98US-0088670P.  
 XX PR 11-JUN-1998; 98US-0089138P.  
 PA (BRAN/ ) BRANCH A. D.  
 PA (WALE/ ) WALENSKI J. L.  
 PA (STUM/ ) STUMP D. D.  
 PA Branch AD, Walewski JL, Stump DD;  
 XX DR WPI; 2000-126431/11.  
 PT Novel Hepatitis C virus peptides useful in vaccine compositions, for  
 PT diagnosing HCV infection and as therapeutic agents.  
 XX Example 1; Page 39-41; 50pp; English.  
 PS  
 XX  
 CC The present sequence is a Hepatitis C virus (HCV) protein derived from an  
 CC HCV isolate AAD14853. The novel HCV protein is not encoded by the  
 CC standard HCV open reading frame but by an alternate reading frame (ARF)  
 CC which is +1 or +2 to the standard HCV ORF. The position of the first  
 CC nucleotide of the ARF may vary slightly depending upon the isolate. The  
 CC protein elicits an immune response in patients infected with HCV and are  
 CC produced during HCV infection. The present sequence is used in vaccine  
 CC compositions for preventing HCV infection. It is also used for diagnosing  
 CC HCV infection and as a target for anti-HCV therapy  
 XX  
 SQ Sequence 198 AA;  
 Query Match 68.0%; Score 583; DB 3; Length 198;  
 Best Local Similarity 72.6%; Prod. No. 5.1e-51;  
 Matches 119; Conservative 3; Nismatches 42; Indels 0; Gaps 0;  
 XX  
 Qy 1 ARILNLKEKPTNVTPTVAHRTSSRVAVRSVLVEFTCRAAGALDWYCARRERLPSGRNLVEYD 60  
 Db 1 ARILNLKEKPTNVTPTVAHRTSSRVAVRSVLVEFTCRAAGAPGWVACRLGTLPSGRNLVEGDNLS 60  
 XX  
 Qy 61 VSLSPRLVYGPRAQPGLSPGTLGPSMAMRAAGGRDGSCLPVALGLAGAQTPGVGRATWVR 120  
 Db 61 VSLSPRPADPREGPSGTLGPSMAMRAVGRDPSCPAAGLVLGAGLUTPGGHAITWVR 120  
 XX  
 Qy 121 SSTPLRASPTSWGTTTSSAPPLLEALPGPWRMASGEFWKTA 164  
 Db 121 SSIPSRVASPTSWGTTTSSALXGALPELWMSWKTAXITQ 164  
 XX  
 RESULT 9  
 ABU61868  
 ID ABU61868 standard; protein; 143 AA.  
 XX  
 AC AC ABU61868;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 14-AUG-2003 (first entry)  
 XX  
 DE HCV core protein frameshift protein p17 #5.  
 XX  
 KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
 KW hepatotropic; overlapping open reading frame; p21c; vaccine.  
 XX  
 OS Hepatitis C virus; genotype 1b.  
 XX  
 FH Location/Qualifiers  
 FT Misc-difference 143  
 /label= unknown  
 FT Misc-difference 161  
 /label= unknown  
 FT Misc-difference 184  
 /label= "Corresponds to stop codon"  
 FT

XX	14-DEC-2000;	20000US-00736959.	OS	Hepatitis C virus; genotype 1b.
XX	PP	US2002076415-A1.	XX	
PR	14-DEC-1999;	99US-0170835P.	XX	
XX	XX	20-JUN-2002.	XX	
PA	(OUUJ/)	OU J.	XX	
PA	(XUZZ/)	XU Z.	XX	
PT	Ou J,	Xu Z;	XX	
PT	DR	2003-479366/45.	XX	
XX	XX	Isolated hepatitis C virus (HCV) proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.	XX	
PT	PS	Claim 9; Page 14; 37pp; English.	XX	
XX	XX	The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21c). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)	XX	
XX	XX	Sequence 143 AA;	XX	
Qy	7	KKCTTNTPTAHRTSSSRVAVRSLEFTCRAGALDTCARRELPSGRNLEYDVSLSPR 66	Query Match	Score 548; DB 6; Length 143;
Db	8	QRKPNVTLTAHRTSSSRVAVRSLEFTCRAGAPGNYCARLGRPLSPGRNLYVGDNLSPR 67	Best Local Similarity	63.9%; Pred. No. 1.3e-17;
Qy	67	LVGPRAGPGLSPGTGPSAMRAAGGRDGSCLPVALGLGAPQPGVGRAIWRSIPLR 126	Matches 106;	Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;
Db	68	LAGPRAGPGLSPGTGPSAMRAAGQDGSCHPAGLGLVGAAPMTPGVGRIVWRSIPLR 127	Qy	7 KKCTTNTPTAHRTSSSRVAVRSLEFTCRAGALDTCARRELPSGRNLEYDVSLSPR 66
Qy	127	AASPTSMGWTYRSSAP 141	Db	8 QRKPNVTLTAHRTSSSRVAVRSLEFTCRAGAPGNYCARLGRPLSPGRNLYVGDNLSPR 67
Db	128	AASPTSMGWTYRSSAP 142	Qy	67 LVGPRAGPGLSPGTGPSAMRAAGGRDGSCLPVALGLGAPQPGVGRAIWRSIPLR 126
XX	XX	RESULT 1.0	Db	68 LADPRAGPGLSPGTGPSAMRAAGQDGSCHPAGLGLVGAAPMTPGVGRIVWRSIPLR 127
AC	ABU61872;	ABU61872 standard; protein; 143 AA.	Qy	127 AASPTSMGWTYRSSAP 141
XX	XX	DT 23-OCT-2003 (revised)	Db	128 AASPTSMGWTYRSSAP 142
DT	14-AUG-2003 (first entry)		RESULT 11	
XX	XX	DB ABU61872 standard; protein; 143 AA.	ABU61871	
DB	XX	HCV core protein frameshift protein p17 #9.	ID ABU61871	
XX	XX	KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; viricide; hepatotropic; overlapping open reading frame; p21c; vaccine.	XX	
KW	XX	DT 23-OCT-2003 (revised)	AC ABU61871;	
KW	XX	DT 14-AUG-2003 (first entry)	XX	

DE HCV core protein frameshift protein p17 #8.  
 XX HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
 KW hepatotropic; overlapping open reading frame; p21C; vaccine.  
 XX Hepatitis C virus; genotype 1b.  
 XX US2002076415-A1.  
 XX PD 20-JUN-2002.  
 XX PP 14-DEC-2000; 20000US-00736959.  
 XX PR 14-DEC-1999; 99US-0170835P.  
 XX PA (OUJZ/) OU J.  
 PA (XUZZ/) XU Z.  
 XX PI Ou J., Xu Z;  
 XX DR WPI; 2003-479366/45.  
 XX PT Isolated hepatitis C virus (HCV) proteins formed by expression of an overlapping open reading frame in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.  
 XX PS Claim 9; Page 15; 37pp; English.  
 CC The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21C). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the anti-viral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)  
 XX Sequence 143 AA;  
 SQ Query Match 63.8%; Score 547; DB 6; Length 143;  
 Best Local Similarity 78.5%; Prod. No. 1..e-47;  
 Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
 Qy 7 KCKKINVTPTAHRTSSSERAVRSLSVFTCRAAGALDVCARRIERLPSGRNLIEDVDSLSLSPR 66  
 Db 8 QRKPNVTPTAAHTSSSERAVRSLSVFTCRAAGALDVCARRIERLPSGRNLIEDVDSLSLSPR 67  
 Qy 67 LYOPRAGEGFLSPGTLGPSPMAMBAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 126  
 Db 68 LASPRAAGLSPGTLGPSPMAMBAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 127  
 Qy 127 AASPTNSCTYRSSAP 141  
 Db 128 AASPTNSCTFRSSAP 142  
 Qy 67 LVGPRAGPGLSPGTLGPSPMAMRAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 126  
 Db 68 LAXPRAAGPGLSPGTLGPSPMAMRAWGQDGSCHPAAPGLVGAAPRTGVRVWNSSPLR 127

AC ABU61867;  
 XX DT 23-OCT-2003 (revised)  
 DT 14-AUG-2003 (first entry)  
 XX DE HCV core protein frameshift protein p17 #4.  
 XX KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
 KW hepatotropic; overlapping open reading frame; p21C; vaccine.  
 XX OS Hepatitis C virus; genotype 1b.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 70 /label= UNKNOWN  
 /note= "Any amino acid"  
 XX PA US2002076415-A1.  
 XX PD 20-JUN-2002.  
 XX PP 14-DEC-2000; 20000US-00736959.  
 XX PR 14-DEC-1999; 99US-0170835P.  
 XX PA (OUJZ/) OU J.  
 PA (XUZZ/) XU Z.  
 XX PS Claim 9; Page 14; 37pp; English.  
 PI Ou J., Xu Z;  
 XX DR WPI; 2003-479366/45.  
 XX PT Isolated hepatitis C virus (HCV) Proteins formed by expression of an overlapping open reading frame in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.  
 XX PA (OUJZ/) OU J.  
 PA (XUZZ/) XU Z.  
 XX PS Claim 9; Page 14; 37pp; English.  
 CC The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21C). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the anti-viral compound may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)  
 CC Sequence 143 AA;  
 SQ Query Match 63.0%; Score 540; DB 6; Length 143;  
 Best Local Similarity 77.8%; Prod. No. 8..2e-47;  
 Matches 105; Conservative 6; Mismatches 24; Indels 0; Gaps 0;  
 Qy 7 KCKKINVTPTAHRTSSSERAVRSLSVFTCRAAGALDVCARRIERLPSGRNLIEDVDSLSLSPR 66  
 Db 8 QRKPNVTPTAAHTSSSERAVRSLSVFTCRAAGALDVCARRIERLPSGRNLIEDVDSLSLSPR 67  
 Qy 67 LYOPRAGEGFLSPGTLGPSPMAMBAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 126  
 Db 68 LASPRAAGLSPGTLGPSPMAMBAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 127  
 Qy 127 AASPTNSCTYRSSAP 141  
 Db 128 AASPTNSCTFRSSAP 142  
 Qy 67 LVGPRAGPGLSPGTLGPSPMAMRAAGGRGSCLPVALGIAAGAPOTPGYGRAIWRSSTPLR 126  
 Db 68 LAXPRAAGPGLSPGTLGPSPMAMRAWGQDGSCHPAAPGLVGAAPRTGVRVWNSSPLR 127

RESULT 12  
 ABU61867  
 ID ABU61867 standard; protein; 143 AA.  
 XX



Best Local Similarity 77.8%; Pred. No. 8.2e-47;  
Matches 105; Conservative 5; Mismatches 25; Indels 0; Gaps 0;

CC standardise OS field  
XX  
SQ Sequence 143 AA:

Query Match 62.3%; Score 534; DB 6; Length 143;  
Best Local Similarity 77.0%; Pred. No. 3.4e-46;  
Matches 104; Conservative 25; Mismatches 0; Indels 0; Gaps 0;

Qy 7 KKCTNVTPTVTAHRTSSSRVAVRSVLYEFTCCRAGALDWMVCA  
Db 8 QRKPNVTPTPAHRTLSSRVSLEFTCCRAGAPGWMVCA  
Qy 67 LVDPRAAGPLSPGTIGPSMAMRAAGGRDSCSPLVAGLIGAQPQPGVRA  
Db 68 LAGPAGPGLSPGTIGPSMAMRAANGQDSSCHPAPGLVGA  
Qy 127 AASPTSWGTYRSSAP 141  
Db 128 AASPTSWGTYRSSAP 142

Qy 7 KKCTNVTPTVTAHRTSSSRVAVRSVLYEFTCCRAGALDWMVCA  
Db 8 QRKPNVTPTAHTSSRAVRSVLYEFTCCRAGAPGWMVCA  
Qy 67 LVGPRAGPGLSPGTIGPSMAMRAAGGRDGSCLPVALGLAGA  
Db 68 LASPRAGPGLSPGTIGPSMAMRAVWGQGSCHVAPGLVGA  
PRTGVRVWRSIIPSH 127

Qy 127 AASPTSWGTYRSSAP 141  
Db 128 AASPTWGTFRSSAP 142

RESULT 15

ABU61875

ID ABU61875 standard; protein; 143 AA.

XX ABU61875

AC AC

XX 23-OCT-2003 (revised)

DT 14-AUG-2003 (first entry)

XX

DE HCV core protein frameshift protein p17 #12.

XX

KW hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
KW hepatotropic; overlapping open reading frame; p21c; vaccine.

XX

OS Hepatitis C virus; genotype 1b.

PN US2002076415-A1.

XX

PD 20-JUN-2002.

XX

PP 14-DEC-2000; 2000US-00736959.

XX

PR 14-DEC-1999; 99US-0170835B.

XX

PA (OUJJ/) OU J.  
(XUZZ/) XU Z.

XX

PI Ou J, Xu Z;

XX

DR WPI: 2003-479366/45.

XX

PT Isolated hepatitis C virus (HCV) Proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.

XX

PS Page 15; 37pp; English.

CC The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21c). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequence that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the anti-viral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to

Copyright GenCore version 5.1.6  
(c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:26:38 ; Search time 40 Seconds  
(without alignments)

394.489 Million cell updates/sec

Title: US-10-664-038-1  
Perfect score: 857  
Sequence: 1 ARILNLKKTNTVTPVAHRT.....ALPGPWRMAMSGFWKTTATMQQ 164

Scoring table: BL05062  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79;\*  
1: pix1;\*  
2: pix2;\*  
3: pix3;\*  
4: pix4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	374	43.6	95	2	S44213		core protein - hep hypothetical prote transcription fact transcription fact hypothetical prote ankyrin related protein transcription fact ABC transporter, A transcription fact transcription fact hypothetical prote protein F5J5_15 (1 collagen dpy-2 - C serum response factor immediate-early protein collagen alpha 1(I) hypothetical prote hypothetical prote hypothetical protein period protein. Period protein. Period protein. Period protein receptor dopamine receptor hypothetical prote elastin precursor hypothetical prote hypothetical prote hypothetical prote hypothetical protein (cytosine-5-) - galactokinase (EC collagen alpha 1(I transcription acti
2	97	11.3	363	2	S22542		
3	96	11.2	479	1	S22543		
4	96	11.2	485	1	S22543		
5	93	10.9	105	2	H72733		
6	90	10.5	1016	2	T19006		
7	89	10.4	479	1	A31753		
8	89	10.4	552	2	F5J311		
9	88	5	583	1	S22544		
10	87	5	10.3	583	I47154		
11	87	5	10.2	478	1	D7622	
12	86	10.2	756	2	B86482		
13	85	10.0	1617	2	T34649		
14	85	9.9	360	2	T37285		
15	85	9.9	365	2	A39481		
16	85	9.9	825	1	E7BEXD		
17	85	9.9	1042	1	CGCH1S		
18	84	9.8	143	2	B72699		
19	84	9.8	152	2	T34649		
20	83	9.8	1113	2	T14260		
21	83	9.7	387	1	DYHJD4		
22	83	9.7	299	2	C70753		
23	83	9.7	784	2	A26601		
24	82	9.6	419	2	G70602		
25	82	9.6	538	2	T27156		
26	82	9.6	1573	2	S01845		
27	81	9.5	395	1	KISMG		
28	81	9.5	671	1	CGRT1S		
29	81	9.5	1442	2	T42607		

## ALIGNMENTS

hypothetical prote transcription fact probable MFS trans collagen alpha 1(I) collagen alpha 6(I) probable DNA-direc transcription fact transcription fact transcription fact transcription fact hypothetical prote collagen alpha 1(I) hook-containing pr cytokine receptor probable infB - My probable PPE prote probable relaxase

hypothetical prote

core protein - hepatitis C virus

C;Species: hepatitis C virus

C;Date: 13-Jan-1995 #text\_change 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S44213

R;Fauchet, H.R.

submitted to the EMBL Data Library, April 1994

A;Reference number: S44213

A;Accession: S44213

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-95 <PEU>

A;Cross-references: UNIPROT:Q68874; EMBL:X78950; PIDN:CAA55547.1; PID:9860

RESULT 1

S44213

core protein - hepatitis C virus

C;Species: hepatitis C virus

C;Date: 13-Jan-1995 #text\_change 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S44213

Query Match 43.6%; Score 374; DB 2; Length 95;

Best Local Similarity 78.7%; Pred. No. 1. 6e-25;

Matches 74; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 11 NVTPTVAHRTSSSSRAVRSLVEFTCRAGALDWVCAARRRLPSGRNLLEVDSLSPRLVGP 70

Db 2 NVTPTAAHRTSSSSRAVRSLVEFTCRAGAPGVNCAARPGRPLPSGRNLVEGDNLSPRLASP 61

Qy 71 RAGPGLSGPTLGPSSMAVRAAGGRDGSCLPVALG 104

Db 62 RAGPGLSGPTLGPSSMAVRAVGQGDGSCHPTAPGL 95

RESULT 2

T34931

hypothetical protein SC3F9\_09 SC3F9\_09 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;DDate: 05-Nov-1999 #text\_change 05-Nov-1999 #text\_change 05-Nov-1999

C;Accession: T34931

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1998

A;Reference number: 221562

A;Accession: T34931

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-363 <SEB>

A;Cross-references: UNIPROT:O69949; EMBL:AL023862; PIDN:CAA19632.1; GSPDB:GN00070; SCOEL

C;Genetics:

A;Gene: SCOEDB:SC3F9\_09

Query Match 11.3%; Score 97; DB 2; Length 363;

Best Local Similarity 30.6%; Pred. No. 0. 47;

Matches 44; Conservative 14; Mismatches 40; Indels 46; Gaps 8;

Qy 39 GALDWVCAARRRLPSGRNLLEVDSLSPRLVGPAGRLPSGRNLVEGDNLSPRLASP 111

Db 244 GALTGLVARRRTRLPAGR-----AAAAPGSGSTPAPAVT-----DA 281  
 Qy 98 LPVÄUGLAGA-----PQPPGVGRAI--WVRSIPLRASPTSGTYRSSAPL 142  
 Db 282 LAAALGAYAATAARPYFAHALINPSSPLTRAVGGIRATIPLQALARSGA--SATSL 339  
 Qy 143 LEAIPGPWMASGFW--KTATMQ 164  
 Db 340 LVAALAP---AGRWFAKRSANRK 359

A;Title: Multiple Oct2 isoforms are generated by alternative splicing.  
 A;Reference number: S22539; MUID:91187647; PMID:2011512  
 A;Accession: S22543  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-485 <WIR>  
 A;Cross-references: EMBL:X57936; NID:953481; PIDN:CAA41004.1; PID:953482  
 A;Note: Splice form Oct-2.1  
 A;Accession: S22540  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-62-124-485 <WI3>  
 A;Cross-references: EMBL:X57941; NID:953491; PIDN:CAA41009.1; PID:953492  
 A;Note: Splice form Oct-2.6  
 R;Istoykova, A.S.; Sterrer, S.; Erselius, J.R.; Hatzopoulos, A.K.; Gruiss, P.  
 Neuron 8, 541-558, 1992  
 A;Title: Mini-Oct and Oct-2c: two novel, functionally diverse murine Oct-2 gene products.  
 A;Reference number: JH0596; MUID:92198662; PMID:1550677  
 A;Accession: JH0596  
 A;Molecule type: mRNA  
 A;Residues: 1-62-85-130, 'A', '432, 'P', 434-473 <STO>  
 A;Cross-references: EMBL:57039; NID:953493; PIDN:CAA40369.1; PID:953494  
 A;Note: Splice form Oct-2c  
 R;Stepchenko, A.G.  
 Dokl. Akad. Nauk SSSR 325, 175-178, 1992  
 A;Title: [Interaction of Oct-binding transcription factors with a large series of 'none' genes]  
 A;Reference number: I48698  
 A;Accession: I48698  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 'P', 189-430, 'A', '432, 'P', 434-464, 'W', 466-485 <RES>  
 A;Cross-references: EMBL:57039; NID:953493; PIDN:CAA40369.1; PID:953494  
 C;Comment: This protein is a tissue-specific transcriptional trans-activator.  
 A;Gene: Oct-2  
 A;Map position: I48698  
 A;Map Position: 7  
 A;Map Position: 7  
 C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology <POU>  
 C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulator  
 P;202-269/Domain: POU domain homology <POU>  
 P;298-354/Domain: homeobox homology <HOX>  
 A;Gene: Oct-2  
 A;Map position: 7  
 A;Introns: 221/3; 269/1; 318/3  
 C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology  
 C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulator  
 P;202-269/Domain: POU domain homology <POU>  
 P;298-354/Domain: homeobox homology <HOX>

Query Match 11-2%; Score 96; DB 1; Length 479;  
 Best Local Similarity 25.4%; Pred. No. 0.76; Gaps 6;  
 Matches 34; Conservative 8; Mismatches 28; Indels 64; Gaps 6;

Qy 43 WVCARRER-----LPS-GRNLVDVSLSPRLVPGPAGPGLSP-----78  
 Db 344 WFCNRQKEKRINPCASAPMLPSPEK----PTSTSPHLVTPQGAGTLPSQASSLST 399  
 Qy 79 -----GTLGPNSHAMRAAGRDGSCLPV-----ALGL 104  
 Db 400 VTTLSAVVGTIHPSPRTAGGGGGGRALPLNSIPSPVTPPPATINNSPQGSHSAAGL 459  
 Qy 105 AGAPQTPGVGRAI 118  
 Db 460 SGL--NPSAGPGLW 471

Query Match 11-2%; Score 96; DB 1; Length 485;  
 Best Local Similarity 25.4%; Pred. No. 0.77; Gaps 6;  
 Matches 34; Conservative 8; Mismatches 28; Indels 64; Gaps 6;

Qy 43 WVCARRER-----LPS-GRNLVDVSLSPRLVPGPAGPGLSP-----78  
 Db 350 WFCNRQKEKRINPCASAPMLPSPEK----PTSYSPHLVTPQGAGTLPSQASSLSTT 405  
 Qy 79 -----GTLGPNSHAMRAAGRDGSCLPV-----ALGL 104  
 Db 406 VTLSSAVVGTIHPSPRTAGGGGGGRALPLNSIPSPVTPPPATINNSPQGSHSAAGL 465

Qy 105 AGAPQTPGVGRAI 118  
 Db 466 SGL--NPSAGPGLW 477

## RESULT 4

S22543 transcription factor Oct-2 splice form Oct-2.3 - mouse  
 N;Alternative names: NF-A2 protein; Oct-2 protein  
 N;Contains: transcription factor Oct-2 splice form Oct-2.1; transcription factor Oct-2  
 C;Species: Mus musculus (house mouse)  
 C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C;Accession: S22543; S22541; S22540; JH0596; I48698; S47215  
 R;Istoykova, A.; Annweiler, A.; Zwilling, S.; Oeler, B.  
 Nucleic Acids Res. 19, 43-51, 1991

## RESULT 5

H72733 hypothetical protein APE0408 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C;Accession: H72733  
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatai, K.; Jin-no, K.; Takaiava, H.; Takamiya, M.; Mabuda, S.; Funahashi, T.; Tanaka, T.; Kudooh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyperthermophilic Crenarchaeon, Aeropyrum  
 A;Reference number: A72450; MUID:9931039; PMID:10382966  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-105 <KAW>  
 A;Cross-references: UNIPROT:Q9YF32; DDBJ:AP000059; NID:g5103911; PID:dr1  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APE0408

Query Match 10.9%; Score 93.5; DB 2; Length 105;  
 Best Local Similarity 33.7%; Pred. No. 0.27; Indels 23; Gaps 6;  
 Matches 34; Conservative 11; Mismatches 33; CLE>  
 A;Accession: C31753  
 A;Cross-references: UNIPROT:PO9086; EMBL:X53468; NID:g35128; PID:CAA37562.1; PMID:935124

Query 64 SPRIVG--PRAGELSPCTLGPSMAMRAG-----GRDSSCLPVA-----LGLAG 107  
 Db 5 SIRAVSGTLPRAASP--PSLGLKWNFLSGRGGACIPSLTNILLGTS 59  
 A;Accession: T19006; T22086  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone C06C3  
 A;Accession: T22086  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone F42B8

RESULT 6  
 T19006  
 ankyrin related protein C06C3.1 - Caenorhabditis elegans  
 N;Contains: myosin-light-chain-phosphatase (EC 3.1.3.53)  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T19006; T22086  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone C06C3  
 A;Accession: T219058  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone C06C3  
 A;Accession: T219050  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone C06C3  
 A;Accession: T219510  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone C06C3  
 A;Accession: T22086  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:CG  
 A;Experimental source: clone F42B8

Query Match 10.5%; Score 90; DB 2; Length 1016;  
 Best Local Similarity 22.3%; Pred. No. 5.3; Indels 54; Gaps 6;  
 Matches 39; Conservative 22; Mismatches 60; CLE>  
 A;Accession: C06C3.1  
 A;Map position: 2  
 A;Introns: 27/3; 94/3; 279/3; 352/2; 462/2; 523/3; 569/2; 657/2; 718/3; 766/3; 833/3; 89  
 C;Keywords: Phosphoric monoester hydrolase

Query 6 LKKCTNVTPTVAHRTSSSRVAV-RSLVLFETCCRAGALDWVCAARRLFPSGRNLLEVDSL 64  
 Db 439 LKIGKEISSLRSETTSSRISITSLDGT-----DRSSSGRETSAAEMS- 482  
 A;Accession: C06C3.1  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; PMID:94379032

Query 65 PRVYGPRAAGLSPGTLG-----PSMAMBAAGG-----RD 94  
 Db 483 -----EARSSASTGTSSSSRFTSISITSLDGT-----DRSSSGRETSAAEMS- 536  
 A;Accession: C06C3.1  
 A;Cross-references: UNIPROT:Q17718; EMBL:236719; PIDN:CAA85318.1; PMID:94379032

Query 95 GSCLPVPAVLAGAQTPGIVGRAIWRSRSPSSWVGRVPLSSRSSSTSVRSSTSPVSEIIISPP 149  
 Db 537 SATIPI-VPLSAPPKAVHICSPSSWVGRVPLSSRSSSTSVRSSTSPVSEIIISPP 590  
 A;Accession: GDB:POU2F2; OCT2; OTF2

A; Title: Multiple Oct2 isoforms are generated by alternative splicing.
A; Cross-references: GDB:120255; OMIM:164176
A; Map position: 19pter-19qter
A; Introns: 221/3; 269/1; 318/3
C; Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
P; 298-354/Domain: POU domain homology <POU>
Query Match 10.4%; Score 89; DB 1; Length 479;
Best Local Similarity 23.9%; Pred. No. 3;
Matches 32; Conservative 10; Mismatches 28; Indels 64; Gaps 6;
Qy 43 WVCARRER-----LPS-GRNLLEVDSLSPRLVSPRAGPLSP----- 78
Db 344 WPCNRRQKEKRINPCSAAPLMPSPSK-----PASTSPHMTPQGAGTTLPSQASSSSLST 399
Qy 79 -----GTLGPMSMARRAGRDSCLPV----- 104
Db 400 VTLSSAVGTLHPSRTAGGGGGGGAAAPPNSIFSVTPPPAPATNSTNSPQGSHSARGL 459
Qy 105 AGAATPQGYGRAIN 118
Db 460 SGL--NPSTCPGLW 471
RESULT 8
F75311 ABC transporter, ATP-binding protein - Deinococcus radiodurans (strain R1)
C; Species: Deinococcus radiodurans
C; Database: 03-Dec-1998 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C; Accession: F75311
R. M. White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
A; Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A; Reference number: A753250; MUID:20036896; PMID:10567266
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-552 <WHI>
A; Cross-references: UNIPROT:Q9RSH9; GB:AE002048; GB:AE000513; NID:96459929; PIDN:AAFI168
A; Experimental source: Strain R1
C; Genetics:
A; Map position: 1
Query Match 10.4%; Score 89; DB 2; Length 552;
Best Local Similarity 29.5%; Pred. No. 3.5;
Matches 49; Conservative 10; Mismatches 55; Indels 52; Gaps 9;
Qy 18 HRT\$SSRAVRSLEFTCCRAGLDWVCCARRLEPSGNLEVDVSLSPRLVGPRACPG----- 75
Db 119 HAAA\$AAR2-----SCAREHHD-VCARR----TGSAFRRAAYEE-----RGGACRR 162
RESULT 10
I47154
Qy 76 -LPGTLCPSMAMARRAGRDGSCLPVAGLAGAPOTPGVGRATW-----VRSSIP 124
Db 163 GVLHSRSLGTDRARRAGRRGGR-TPAAG-SDARTTVEGRDYTDDRGALAGARLSHP 220
Qy 125 LRAASPTSGTYSSA-----PPEALFGP 149
Db 221 LARAPPDGAPRDPDARAAARRIHRMPERSGPRRRRAAEGP 266
RESULT 9
S22544
Transcription factor Oct-2, splice form Oct-2.5 - mouse
N; Alternative names: NPFA2; Oct-2; transcription factor Oct-2b
C; Species: Mus musculus (house mouse)
C; Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C; Accession: S22544; M50085; S03240
R. Wirth, T. Prives, A.; Annweiler, A.; Zwilling, S.; Oeler, B.
A; Nucleic Acids Res. 19, 43-51, 1991
A; Title: Multiple Oct2 isoforms are generated by alternative splicing.
A; Cross-references: UNIPROT:Q00196; EMBL:X57940; NID:953498; PIDN:CAA41008; PID:953499
R. Hatzopoulos, A.K.; Stoykova, A.S.; Erselius, J.R.; Gouling, M.; Neuman, T.; Gruss, P.
Development 109, 349-362, 1990
A; Title: Structure and expression of the mouse Oct2a and Oct2b, two differentially spliced POU domain homologs.
A; Reference number: A60085; MUID:90382251; PMID:1976089
A; Status: not compared with conceptual translation
A; Molecule type: mRNA
A; Residues: 1-583 <WIR>
A; Cross-references: UNIPROT:Q00196; EMBL:X57940; NID:953498; PIDN:CAA41008; PID:953499
R. Hatzopoulos, A.K.; Stoykova, A.S.; Erselius, J.R.; Gouling, M.; Neuman, T.; Gruss, P.
Development 109, 349-362, 1990
A; Title: Structure and expression of the mouse Oct2a and Oct2b, two differentially spliced POU domain homologs.
A; Reference number: A60085; MUID:90382251; PMID:1976089
A; Status: not compared with conceptual translation
A; Molecule type: mRNA
A; Residues: 1-408 /'A', 410, 'F', 412-583 <HAT>
A; Cross-references: EMBL:X33654; NID:9288173; PIDN:CAA37702; PID:9288174
R. Goldsborough, A.; Ashworth, A.; Williamson, K.
Nucleic Acids Res. 18, 1631, 1990
A; Title: Cloning and sequencing of POU-boxes expressed in mouse testis.
A; Reference number: S09237; MUID:90221898; PMID:1970171
A; Accession: S09237
A; Cross-references: EMBL:X51961; NID:953495; PIDN:CAA36220; PID:953496
A; Accession: S09240
A; Cross-references: EMBL:X51961; NID:953495; PIDN:CAA36220; PID:953496
C; Comment: This protein is a tissue-specific transcriptional trans-activator.
C; Genetics:
A; Gene: Oct-2
A; Map Position: 7
C; Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
C; Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
P; 186-233/Domain: POU domain homology <POU>
F; 282-338/Domain: homeobox homology <HOX>
Query Match 10.3%; Score 88.5; DB 1; Length 583;
Best Local Similarity 22.9%; Pred. No. 4.1; Mismatches 41; Indels 103; Gaps 9;
Matches 47; Conservative 14; Mismatches 41; Indels 103; Gaps 9;
Qy 43 WVCARRER-----LPS-GRNLLEVDSLSPRLVSPRAGPLSP----- 78
Db 328 WFCNTRQKEKRINPCSAAPLMPSPGK-----PTSYSPHLVTPQGAGTLPLSQASSSSLST 383
Qy 79 -----GTLGSMANAAGGDRGDSCLPV----- 104
Db 384 VTISSAVGTLHSRTAGGGGGGRALPLNSLPSVTPPPPATTNSTNSPOGGSNSAIGL 443
Qy 105 AGAQQTGPG---VG----- 130
Db 444 SGUNPSAGSTMVGLSSGUSPALMSNNPLATQALASGTTPLTSLDGSGNVLGAAGAAP 503
RESULT 10
I47154
Transcription factor Oct-2 - pig
N; Alternative names: NPFA2; Oct-2
C; Species: Sus scrofa domestica (domestic pig)
C; Date: 21-Feb-1997 #sequence_revision 21-Feb-1997
C; Accession: I47154
R. Triggle, C.K.; Helm, J.; Rothschild, M.F.
Anim. Genet. 25, 141-145, 1994
A; Title: Cloning, sequencing and restriction fragment length polymorphism analysis of a
A; Reference number: I47154; MUID:95030552; PMID:7943947
A; Accession: I47154
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
A; Residues: 1-478 <TUG>
A; Cross-references: UNIPROT:Q29013; EMBL:U00794; NID:9451313; PIDN:AAA80148; PID:94511
C; Comment: This protein is a tissue-specific transcriptional trans-activator.
C; Genetics:
A; Gene: OCT2
C; Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulator  
 P;202-269;Domain: POU domain homology <POU>  
 P;298-354;Domain: homeobox homology <Hox>

Query Match 10.2%; Score 87.5; DB 1; Length 478;  
 Best Local Similarity 26.7%; Pred. No. 4.1;  
 Matches 44; Conservative 9; Mismatches 37; Indels 75; Gaps 10;

Qy 43 WVCARRIER -----LPS-GRNLLEVDSLSSRLVPGPRAGPSP----- 78  
 Db 344 WFCNRBQEKEKRINPCSAAPMLPSPGK--PASYSPHLVTBQGGTTLPSQASSSLST 399

Qy 79 -----GTLGPSPMAMRAAQRGRDGSCLPVALGLAGAPOTPGVGRAIWRSSPIPLRAASP 130  
 Db 400 VTTLSSAVGTHPS--RTAGG-----GAACGGAAPL-----NSIP--SVTP 437

Qy 131 TSWGTYRSAPL-----LEALPGPWRMASGFWKATMQ 163  
 Db 438 PPPATTNSTNPQGSHSATGLSGLNPSTGP-----GLNNPAPYQ 477

RESULT 11

D75522 hypothetical protein - *Deinococcus radiodurans* (strain R1)  
 C;Species: *Deinococcus radiodurans*  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: D75522  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vaishnav, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A;Reference number: A75280; MUID:20036896; PMID:10567266  
 A;Accession: D75522  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-756 <WHI>  
 A;Cross-references: UNIPROT:Q9RZS5 ; GB:AE001826; NID:956460827; PID:9646  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DRB0039  
 A;Map position: megaplasmid  
 A;Genome: Plasmid  
 A;Note: plasmid MP1

Query Match 10.2%; Score 87; DB 2; Length 756;  
 Best Local Similarity 25.6%; Pred. No. 7.2;  
 Matches 46; Conservative 14; Mismatches 62; Indels 58; Gaps 9;

Qy 10 TNVTPVVAHTSSSRVAVRSVLEFICRCAGALDWCVARRERLPSGRNLLEVDSLSSRLVGS 69  
 Db 2865 SNATADLIRRVEΤΟΡΕΓLΙRΚ---TVSKVGPDW-ΑΡΓSAVTSCDELYTLSTVNPYVQ 339

Qy 70 PRAG-PGLSPOTLGSMAAAGGRGDSCLPVALGLAGAQ-----PGVGRAIWR 120  
 Db 340 PLAGQVQDLPPLPAGTEFVSAASDGG-----ALLGAASAAQVANTLGDPAGATRLTR 392

Qy 121 -----SSPI-----RAASPTSWGTYRSSAPLL-----EALPG 148  
 Db 393 VRVGRDVRDDQELRNVFEITSSELPAPLHSNAASAVWNT--APLISKLDRRDAAPG 448

RESULT 12

B86483 protein F5J5.15 (imported) - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C;Accession: B86483  
 R/Theology, A.; Ecker, J.R.; Palm, C.J.; Federer, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Deward, K.; anson, N.P.; Hughes, B.; Huijar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C;A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwarz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talion, A;Authors: Salzberg, S.L.; Schwarz, J.R.; Shinn, P.; Southwick, A.M.;  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: B86483;  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1617 <STOP>  
 A;Cross-references: UNIPROT:Q9SKV5 ; GB:AE005172; NID:96599587; PIDN:AAF18642.1; GSPDB:G;  
 C;Genetics:  
 A;Gene: F5J5.15  
 A;Map position: 1

Query Match 10.0%; Score 86; DB 2; Length 1617;  
 Best Local Similarity 26.8%; Pred. No. 19;  
 Matches 34; Conservative 16; Mismatches 55; Indels 22; Gaps 4;

Qy 54 GRNLEVDVDSLSPRL-----VGPRAGPGL-----SPGTIGPSMAMRAAAGRGRDGSCLPVVA 101  
 Db 82 GRSATIEQSYSGSOLFEVPGVGDGLGADAGATSGVAGGAGVSYGHAVGAKDPVGLVGA 141

Qy 102 LGLAGAPQPQGVGRAIWR-----SSPLRANAPTSGYRSAPLALPGFWRMAS 154  
 Db 142 --AGAQVPEVGLAGLRLQOLLERLPGVVPVHAPVAPRVAEVQRAAVBEEVPSYLRMME 198

Qy 155 GFWKAT 161  
 Db 199 QLQRGT 205

RESULT 13

T37285 collagen dpy-2 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
 C;Accession: T37285  
 R;Levy, A.D.; Yang, J.; Kramer, J.M.  
 Mol. Biol. Cell 4, 803-817, 1993  
 A;Title: Molecular and genetic analyses of the *Caenorhabditis elegans* dpy-2 and dpy-10  
 A;Reference number: Z21666; MUID:94060446; PMID:8241567  
 A;Accession: T37285  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Introns: 18/3; 68/3; 118/1; 325/3  
 A;Residues: 1-360 <LEV>  
 A;Cross-references: EMBL:Li12706; NID:9289665; PIDN:AAA17398.1; PIDN:9467810  
 A;Note: a variety of molecular defects in these collagens can result in severe morpholo  
 C;Genetics:  
 A;Gene: dpy-2  
 A;Introns: 18/3; 68/3; 118/1; 325/3  
 C;Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 9.9%; Score 85; DB 2; Length 360;  
 Best Local Similarity 24.4%; Pred. No. 5;  
 Matches 47; Conservative 10; Mismatches 56; Indels 80; Gaps 9;

Qy 32 EFTVCRAGLDW---VCARRERLPSR----- 65  
 Db 63 DFEFOQASANDLTETMMSYREGILGRNTVRAAGYGHYNPSMLADSQFQECBASCIP 122  
 Qy 66 RLVSPRAGPGLS-----PGTIGPSMAMRAAAGRGRDGS-----CLPVVALGLAGA 107  
 Db 123 GERSPGSDGSLPGLPAGRPD-----GAAPRPGTIPNASCIPERFEPPCPQGPAGV 178

Qy 108 PQTPGV-----GRAIWRSIPLRAASPTSWGTY--RASAPL----- 144  
 Db 179 PGHFGPDPGEYGIQGP-----GSDGMPGKGPDPGLAGPIGPQGSSGPIEDKGRTPEAHYI 236

Qy 145 -----ALPGPW 150  
 Db 237 PGFFGDSGLPGPW 249

RESULT 14

A39481 serum response factor-related protein 2 - human  
N:Alternate names: myocyte-specific enhancer factor xnef2; RSRF2  
C:Species: Homo sapiens (man)  
C:Date: 03-Aug-1992 #sequence\_revision 03-Aug-1992 #text\_change 09-Jul-2004  
C:Accession: A39481; S25832; S24468  
R;Pollock, R.; Treisman, R.  
Genes Dev. 5, 2327-2341, 1991  
A;Title: Human SRF-related proteins: DNA-binding properties and potential regulatory targets  
A;Reference number: A39481; MUID:92084105; PMID:1748287  
A;Accession: A39481  
A;Molecule type: mRNA  
A;Residues: 1-165 <POL>  
A;Cross-references: UNIPROT:002080; EMBL:X63380; NID:936168; PID:CAA44988; PID:936169  
R;Yu, Y.T.; Breitbart, R.E.; Smoot, L.B.; Lee, Y.; Mandavi, V.; Nadal-Ginard, B.  
Genes Dev. 6, 1783-1798, 1992  
A;Title: Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted  
A;Reference number: S25830; MUID:92387551; PMID:1516833  
A;Accession: S25832  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-365 <YUY>  
A;Cross-references: EMBL:X68502; NID:G37991; PID:CAA48515.1; PID:937992  
C;Keywords: DNA binding; homodimer; transcription factor  
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match, Score 9.9%; Best Local Similarity 9.9%; Pred. No. 5.1; Length 365; Matches 36; Conservative 10; Mismatches 47; Indels 50; Gaps 7;

Qy 47 RERBDPSGRNLEVDVSLSPVGVPPGPGPUSPGLPQGPMAMRAAGGRDSCLPVALG-L 104  
Db 207 RRSIDLPGG-----LAGRGG-----INTSSRLS-----GLQNPCSTATPGPL 245

Qy 105 AGAAPOTPG---VGRAIWWR-----SSIPRLRAASPTSNQTYRSSAAPIA 145  
Db 246 GSFPPLPGPPVGABAWARRVVPQAPAPRPPQPSASSLASSLRLPPGAPATFLRSPPIPOS 305

Qy 146 LPGPKRMAASGF-----WKT 160  
Db 306 SPGPHQSLCGLGPPCAGCPNPTA 328

## RESULT 15

EDBEXD

1immediate-early protein RL2 - human herpesvirus 2 (strain HG52)

N:Alternate names: RL2 protein  
C:Species: human herpesvirus 2  
A;Note: host Homo sapiens (man)  
C:Accession: JQ1501  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
R;McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.  
J. Gen. Virol. 72, 3057-3075, 1991  
A;Title: Comparative sequence analysis of the long repeat regions and adjoining parts of  
A;Reference number: JQ1194; MUID:9211549; PMID:166297  
A;Accession: JQ1501  
A;Molecule type: DNA  
A;Residues: 1-825 <MCG>  
A;Cross-references: UNIPROT:B28284; GB:D10471; DDBJ:D01128; NID:g221784; PIDN:BA23427.1  
C;Genetics:  
A;Gene: RL2  
A;Introns: 25/3; 252/1  
C;SuperFamily: herpesvirus immediate-early protein IRI10; RING finger homology  
C;Keywords: DNA binding; immediate-early protein; tandem repeat; transcription regulation  
F;122-172/Domain: RING finger homology <RING>  
F;126-156/Region: zinc finger C3HC4 motif  
F;589-623/Region: 5-residue repeats (A-S-S-S-S)

Query Match, Score 9.9%; Best Local Similarity 27.0%; Pred. No. 12; Mismatches 63; Indels 34; Gaps 5;

ALIGNMENTS										
BLOSUM62										
dapop 10.0 , Gapext 0.5										
searched: 1612378 seqs, 512079187 residues										
total number of hits satisfying chosen parameters: 1612378										
minimum DB seq length: 0										
maximum DB seq length: 2000000000										
post-processing: Minimum Match 0%										
Maximum Match 100%										
Listing First 45 summaries										
database : UniProt_03:*										
1: uniprot_sprot:*										
2: uniprot_trembl:*										
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.										
SUMMARIES										
result No.										
Score										
Query										
Match										
Length										
DB										
ID										
Description										
1										
577	67.3	125	2	Q66VR7	Q66vr7 hepatitis C	Q66vr7 hepatitis C	Q66vr7 hepatitis C	Q66vr7 hepatitis C	Q66vr7 hepatitis C	
2	65.0	125	2	Q66VM9	Q66vm9 hepatitis C	Q66vm9 hepatitis C	Q66vm9 hepatitis C	Q66vm9 hepatitis C	Q66vm9 hepatitis C	
3	55.6	64.9	2	Q66VK1	Q66vk1 hepatitis C	Q66vk1 hepatitis C	Q66vk1 hepatitis C	Q66vk1 hepatitis C	Q66vk1 hepatitis C	
4	432.5	50.5	126	2	Q66VN1	Q66vn1 hepatitis C				
5	403	47.0	122	2	Q66VH8	Q66vh8 hepatitis C				
6	374	43.6	95	2	Q66VH74	Q66vh74 hepatitis C				
7	350.5	40.9	134	2	Q70GD2	Q70gd2 hepatitis C				
8	346	40.4	102	2	Q66361	Q66361 hepatitis C				
9	336	39.2	102	2	Q68358	Q68358 hepatitis C				
10	331.5	38.7	102	2	Q63356	Q63356 hepatitis C				
11	319	37.2	102	2	Q68359	Q68359 hepatitis C				
12	305	35.6	100	2	Q68360	Q68360 hepatitis C				
13	301	35.1	71	2	Q9JAW2	Q9jaw2 hepatitis C				
14	296	34.5	71	2	Q9JAW1	Q9jaw1 hepatitis C				
15	287	33.5	102	2	Q68365	Q68365 hepatitis C				
16	285	33.3	102	2	Q63366	Q63366 hepatitis C				
17	262	30.6	102	2	Q68363	Q68363 hepatitis C				
18	262	30.6	102	2	Q68368	Q68368 hepatitis C				
19	254	29.6	102	2	Q68367	Q68367 hepatitis C				
20	225	26.3	80	2	Q81289	Q81289 hepatitis C				
21	220.5	25.7	119	2	Q68686	Q68686 hepatitis C				
22	148	17.3	53	2	Q9JRM9	Q9jrm9 hepatitis C				
23	142.5	16.6	125	2	Q66VTO	Q66vto hepatitis C				
24	142.5	16.5	125	2	Q66VVI	Q66vvi hepatitis C				
25	142.5	16.6	125	2	Q66VV3	Q66vv3 hepatitis C				
26	137.5	16.0	125	2	Q66VP7	Q66vp7 hepatitis C				
27	137.5	16.0	125	2	Q66VQ5	Q66vq5 hepatitis C				
28	123.5	14.4	210	2	Q66VYL1	Q66vyl1 hepatitis C				
29	109.5	12.8	108	2	P87760	P87760 hepatitis C				
30	109.5	12.8	108	2	P87761	P87761 hepatitis C				
31	102.5	12.0	514	1	IWA_ANTELLA	IWA_ANTELLA hepatitis C				



FT	NON-TER	122	122 AA;	122 MW;	PB5C7A7F76AE29AC CRC64;	OX	NCBI_TaxID=41856;
SQ	SEQUENCE	122 AA;				RN	SEQUENCE FROM N.A.
Query Match	47.0%	Score 403;	DB 2;	Length 122;		RX	PubMed:14718617; DOI=10.1099/vir.0.19472-0;
Best Local Similarity	68.6%	Pred. No. 1.9e-24;				RA	Cahn R., Casane D., Vasquez S., Garcia L., Chunga A., Romero H.,
Matches	83;	Conservative	9;	Mismatches 29;	Indels 0;	RA	Kahn B., Cristina J.,
Qy	6 LKKKTNVTPVTAHRTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 65					RA	"Genetic analysis of Hepatitis C viruses in Peruvian patients.";
Db	1 LKEKEPKETPSVVAHRTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 60					DR	EMBL: AJ582128; CABE46584.1; -
Qy	66 RLVRGRAGPGLSPGTLGPSPMAMRAGGRDSCLPVLAGLAGAPOPVGVRAIWKRSSIFL 125					DR	GO_0019028; C: viral capsid; IEA,
Db	61 RRVGAKAGPGLSPGTLGPSPMTRGGGGSCPHAPVPLGAQMTPGEPAIWVKKSSIFL 120					DR	DR InterPro; IPR02322; HCV capsid.
Qy	126 R 126					DR	Pfam: PF01543; HCV capsid; 1.
Db	121 R 121					FT	NON-TER 134 AA; 14214 MW; 85D9075FB32CABAC CRC64;
Qy	Q68874	PRELIMINARY;	PRT;	95 AA.		SQ	SEQUENCE 134 AA;
AC	Q68874;					Query Match	40.9%; Score 350.5;
DT	01-NOV-1996	(TREMBLrel. 01, Created)				Best Local Similarity	85.4%; Pred. No. 2.9e-20;
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)				Matches	70; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)				Qy	52 PSGRNLVDVSLSPRLVPGPRAFGPLSPCTLGPSMAMRAAVGRDGSCLPVALGLAGAPOTP 111
DE	Core protein (Fragment).					Db	58 PRGRR----QTPPRHYPGRAGPGLSPCTLGPSMAMRAAVGRDGSCLPVALGLAGAPOTP 112
OS	Hepatitis C virus.					Db	112 GYGRAIWKRSSIPLRASPTSW 133
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					Db	113 GYGRAIWKRSSIPLRASPTSW 134
OC	Hepacivirus.					RESULT 8	
NCBI_TaxID	11103;					Q688361	
RN	[1]					ID	Q688361
RP	SEQUENCE FROM N.A.					AC	Q688361
RX	MEDLINE:9520681; PubMed:7751366;					PRELIMINARY;	PRT;
RA	Feucht H.H., Zoellner B., Poliakwa S., Laufs R., Laufs R.,					DT	01-NOV-1996 (TREMBLrel. 01, Created)
RT	"Study on reliability of commercially available hepatitis C virus antibody tests."					DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
RT	antibody tests."					DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
RL	Core protein (Fragment).					DE	Core protein (Fragment).
RL	Hepatitis C virus.					OS	Hepatitis C virus.
RL	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
RN	[2]					OC	Hepacivirus.
RP	SEQUENCE FROM N.A.					NCBI_TaxID	11103;
RA	Feucht H.;					RN	[1]
RL	Submitted (APR-1994) to the EMBL/GenBank/DBBJ databases.					RP	SEQUENCE FROM N.A.
DR	EMBL: X88950; CAA5547.1; -.					RC	STRAIN=NEUTK3;
DR	PIR: S4213; S4423.1					RX	PubMed:7595353;
FT	NON-TER 1	95	95 AA;	9702 MW;	647C80587C6FB92F CRC64;	RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
SQ	SEQUENCE 95 AA;					RA	The Int'l. HCV Collaborative Study Group;
FT	NON-TER 1	95				RT	"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification.";
FT	NON-TER 1	95				RT	J. Gen. Virol. 76:2433-2507(1995).
FT	NON-TER 1	95				RL	Feucht H.;
FT	NON-TER 1	95				RA	RA
FT	NON-TER 1	95				DR	Mellor J., U31237; AAA75051.1; -.
FT	NON-TER 1	95				FT	FT
FT	NON-TER 1	95				NON-TER	102 102
FT	NON-TER 1	95				FT	102 102
FT	NON-TER 1	95				SEQUENCE	102 AA; 10446 MW; D3C18355326B429D CRC64;
FT	NON-TER 1	95				Query Match	40.4%; Score 346;
FT	NON-TER 1	95				Best Local Similarity	69.9%; Pred. No. 5.1e-20;
FT	NON-TER 1	95				Matches	72; Conservative 6; Mismatches 23; Indels 2; Gaps 2;
Qy	11 NVTPTVTAHRTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 70					Qy	9 KTNVTPVTAHRTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 68
Db	2 NVTPTAHTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 61					Db	1 KTNVTPVTAHRTSSSRVAVRSLVEFTCRAGALDWVARRRLPSGRNLVDVSLSP 60
Qy	71 RAGPGLSPCTLGPSMAMRAGGRDSCLPVALGL 104					Qy	69 GPRAGPGLSPCTLGPSMAMRAGGRDSCLPV-ALGLAGAPOT 110
Db	62 RAGPGLSPCTLGPSMAMRGGDSCLPVALGL 95					Db	61 SPRAGPSPCTLGPSMAMRGGDSCLPVALGLAPMT 102
RESULT 7						RESULT 9	
07GD2	PRELIMINARY;					Q68358	
AC	Q70GD2;					ID	Q68358
DT	05-JUL-2004	(TREMBLrel. 27, Created)				PRELIMINARY;	PRT;
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)				AC	Q68358
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)				OC	OC
DE	Core protein (Fragment).					OC	OC
OS	Hepatitis C virus type 1.					OC	OC
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					OC	OC

DT	01-NOV-1996	(TREMBLrel. 01, Created)
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DB		Core protein (Fragment).
OS		Hepatitis C viruses.
OC		ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC		Hepadviridae.
OX		NCBI_TaxID=11103;
RN		[1]
RP		SEQUENCE FROM N.A.
RC		STRAIN=EBUH3455;
RC		MEDLINE=96030859; PubMed=7595353;
RX		Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA		The Int'l HCV Collaborative Study Group;
RA		"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification";
RT		J. Gen. Virol. 76:2493-2507(1995).
RL		EMBL: U31254; AAA75058.1; -.
DR		NON_TER 1
FT		NON_TER 102 AA, 102 MW, E3B0611FF5880B35 CRC64;
PT		SEQUENCE 102 AA;
PT		Score 39.4%; Best Local Similarity 69.7%; Pred. No. 3.1e-19; Matches 69; Conservative 2; Mismatches 28; Indels 0; Gaps 0;
Qy		9 KTNVTTVAHRTSSRAVRSVLVEFTCRAGALDWVCCRERLPSGRNLLEVDSLSPRLV 68
Db		1 KPNVTIAPTTSSRAVRSLAETFCRAGALDWVCCRERLPSGRNLLEVDSLSPXA 60
Qy		69 GPRAGPGLSPGTLGPSPMAMRAAGGGRGSCLPVALGLAGA 107
Db		61 GPRAGPGLSLGILGPPTEMRVAGQQGSCHPAVLGLGA 99
RESULT 10		
Q68356		PRELIMINARY;
Q68356		PRT;
AC		102 AA.
DT	01-NOV-1996	(TREMBLrel. 01, Created)
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE		Core protein (Fragment).
DE		Hepadviridae.
OS		ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC		Hepadviridae.
OC		NCBI_TaxID=11103;
RN		[1]
RP		SEQUENCE FROM N.A.
RC		STRAIN=EUUK43;
RC		MEDLINE=96030859; PubMed=7595353;
RA		Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA		The Int'l HCV Collaborative Study Group;
RA		"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification";
RT		J. Gen. Virol. 76:2493-2507(1995).
RL		EMBL: U31254; AAA75056.1; -.
DR		NON_TER 1
FT		NON_TER 102 AA, 102 MW, 3088BC3CDB0F80B396 CRC64;
PT		SEQUENCE 102 AA;
PT		Score 38.7%; Best Local Similarity 68.4%; Pred. No. 7.1e-19; Matches 67; Conservative 7; Mismatches 21; Indels 3; Gaps 1;
Qy		9 KTNVTTVAHRTSSRAVRSVLVEFTCRAGALDWVCCRERLPSGRNLLEVDSLSPRLV 68
Db		1 KPKETIAPTTSSRAVRSLAETFCRAGALDWVCCRERLPSGRNLLEVDSLSPXA 60
Qy		69 GPRAGPGLSPGTLGPSPMAMRAAGGGRGSCLPVALGLGA --LPVAG 103

Db	RESULT 11	Q68359	PRELIMINARY;	PRT;	102 AA.
	ID	Q68359			
	AC	Q68359;			
	DT	01-NOV-1996 (TREMBLrel. 01, Created)			
	DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
	DB	Core Protein (Fragment).			
	OS	Hepatitis C virus.			
	OC	Viruses; sRNA positive-strand viruses, no DNA stat.			
	OX	NCBI_TaxID=11103;			
	RN				
	RP	SEQUENCE FROM N.A.			
	RC	STRAIN=EUGE15;			
	RX	MEDLINE=9603059; PubMed=7595353;			
	RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simm			
	RA	The Int'l. HCV Collaborative Study Group;			
	RT	"Investigation of the pattern of hepatitis C virus			
	RT	in different geographical regions: implications for			
	RT	classification.";			
	RL	J. Gen. Virol. 76:2493-2507(1995).			
	DR	EMBL: U31255; AAA7059..1; -.			
	FT	NON-TER 1			
	FT	NON-TER 102			
	SQ	SEQUENCE 102 AA; 10607 MW; C6F24F1A392EEB87 CRR			
Qy		9 KTNNTPTVAHRTSSRRAVRSLSVFTCCRAGALDWVCARREI	37.2%; Score 319; DB 2; Len 67;		
Db		1 KPNKPTPTTAPWTTSSRAVRSLSAFTCCRAAGALDWVCARLG	68.4%; Pred. No. 6.9e-18; Mismatches 2; Conservative		
Qy		69 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 10			
Db		69 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 10			
		61 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 98			
Qy		9 KTNNTPTVAHRTSSRRAVRSLSVFTCCRAGALDWVCARREI	37.2%; Score 319; DB 2; Len 67;		
Db		1 KPNKPTPTTAPWTTSSRAVRSLSAFTCCRAAGALDWVCARLG	68.4%; Pred. No. 6.9e-18; Mismatches 2; Conservative		
Db	RESULT 12	Q68350	PRELIMINARY;	PRT;	100 AA.
	ID	Q68350			
	AC	Q68350;			
	DT	01-NOV-1996 (TREMBLrel. 01, Created)			
	DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
	DB	Core Protein (Fragment).			
	OS	Hepatitis C virus.			
	OC	Viruses; sRNA positive-strand viruses, no DNA stat.			
	OX	NCBI_TaxID=11103;			
	RN				
	RP	SEQUENCE FROM N.A.			
	RC	STRAIN=EUGE28;			
	RX	MEDLINE=9603059; PubMed=7595353;			
	RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simm			
	RA	The Int'l. HCV Collaborative Study Group;			
	RT	"Investigation of the pattern of hepatitis C virus			
	RT	in different geographical regions: implications for			
	RT	classification.";			
	RL	J. Gen. Virol. 76:2493-2507(1995).			
	DR	EMBL: U31256; AAA7050..1; -.			
	FT	NON-TER 1			
	FT	NON-TER 100			
	SQ	SEQUENCE 100 AA; 10414 MW; 727D8920E9BB281C CRR	35.6%; Score 305; DB 2; Len 67;		
Qy		9 KTNNTPTVAHRTSSRRAVRSLSVFTCCRAGALDWVCARREI	35.6%; Score 305; DB 2; Len 67;		
Db		1 KPNKPTPTTAPWTTSSRAVRSLSAFTCCRAAGALDWVCARLG	67.0%; Pred. No. 8.6e-17; Mismatches 2; Conservative		
Qy		69 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 10			
Db		69 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 10			
		61 GPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVALGLAG 98			



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

MM protein - protein search, using SW model  
run on: August 25, 2005, 02:39:56 ; Search time 22 Seconds  
556.475 Million cell updates/sec

title: US-10-664-038-1  
perfect score: 857  
sequence: 1 ARILNLKKTNVTPVWAHRT.....ALFGPWRMASGFWKATMQQ 164

scoring table: BLOSUM62  
Gapext 10.0 , Gapext 0.5

searched: 513545 seqs, 74649064 residues

total number of hits satisfying chosen parameters: 513545

minimum DB seq length: 0  
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
Listing First 45 summaries

database : Issued Parents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB\_pep:  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB\_pep:  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB\_pep:  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB\_pep:  
5: /cgn2\_6/prodata/1/iaa/PECTUS\_COMB\_pep:  
6: /cgn2\_6/prodata/1/iaa/backf1181.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	100.0	164	US-09-644-987-1	Sequence 1, Appli
2	324	37.8	115	4	Sequence 148, App
3	100	11.7	236	4	Sequence 27618, A
4	99.5	11.6	263	4	Sequence 16906, A
5	98	11.4	162	4	Sequence 18201, A
6	98	11.4	202	4	Sequence 24790, A
7	97	11.3	231	4	Sequence 22801, A
8	96	11.2	709	4	Sequence 27305, A
9	95.5	11.1	328	4	Sequence 17729, A
10	95	11.1	371	4	Sequence 25006, A
11	93	10.9	213	4	Sequence 24250, A
12	92.5	10.8	254	4	Sequence 23311, A
13	91.5	10.7	461	4	Sequence 24717, A
14	91	10.6	253	4	Sequence 29631, A
15	91	10.6	1027	4	Sequence 26216, A
16	90	10.5	220	4	Sequence 16739, A
17	90	10.5	242	4	Sequence 31425, A
18	89.5	10.4	138	4	Sequence 26931, A
19	89.5	10.4	1209	4	Sequence 25844, A
20	89.5	10.4	1225	4	Sequence 25018, A
21	89	10.4	348	4	Sequence 25605, A
22	89	10.4	425	4	Sequence 20467, A
23	89	10.4	470	4	Sequence 9859, Ap
24	88.5	10.3	171	4	Sequence 23216, A
25	88	10.3	146	4	Sequence 27352, A

RESULT 2  
US-09-644-987-1  
; Sequence 1, Application US/09644987  
; Patent No. 6803214  
; GENERAL INFORMATION:  
; APPLICANT: MAVROMATA, PENELOPE  
; APPLICANT: VARALIOTI, AGgelits  
; APPLICANT: GEORGPOULOU, URANIA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH C VIRUS CORE GENE  
; AND/OR OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE  
; TITLE OF INVENTION: PRODUCTS  
; FILE REFERENCE: 03495-0194-00000  
; CURRENT APPLICATION NUMBER: US/09/644-987  
; CURRENT FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/151,074  
; PRIOR FILING DATE: 1999-08-27  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 1  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-644-987-1

ALIGNMENTS

CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 148  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-09-878-281A-148

Query Match 37.8%; Score 324; DB 4; Length 115;  
Best Local Similarity 63.0%; Pred. No. 2.e-26;  
Matches 68; Conservative 4; Mismatches 34; Indels 2; Gaps 1;

Qy 7 KCTTNVTPVAHRSRVAVRSVEFTCRAGLDWVCCRRLPSGRNLLEVDSLSPR 66  
Db 10 KTRNTNP-GHRTLSSQANVRSVEFTCRAGAPSVTCVQCARLPSGRNLAVGANPSPG 67

Qy 67 LVGPRAGPGLSPGTLPSPMAMRAGGRDQSCLPVALGLAGAPQTPGVC 114  
Db 68 RAERAGPGLSPGTLPSPGTYMGRAGQGSCPAAALARQANTPGAG 115

RESULT 3  
US-09-252-991A-27618  
; Sequence 27618; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27618  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27618

Query Match 11.7%; Score 100; DB 4; Length 236;  
Best Local Similarity 28.3%; Pred. No. 0.015;  
Matches 39; Conservative 13; Mismatches 52; Indels 34; Gaps 8;

Qy 34 TCCAGAGLDWVCCRRLPSGRNLLEVDSLSPRVLGPRAGPGLSPGTLPSPMAMRAAGGR 93  
Db 46 TCCRSPGPGACRRSRRTS---TVAATTPTQARRSPPPG--PGRSRSPVRPTSPS--- 96

Qy 94 DGSSLVPLVAGLAGAPQTPCVRGAIWRSIPLRAS---PTSGNTYRSSA-----PLLE 144  
Db 97 -----ANASPNSP-TARARYFR-SVARRATTPTTWAASSTMPSPPRBSST 143

Qy 145 ALPQPWNRASGFEWTKTMM 162  
Db 144 GAPPQWRSST---STTIM 158

RESULT 4  
US-09-252-991A-16906  
; Sequence 16906; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16906

Query Match 11.6%; Score 99.5; DB 4; Length 263;  
Best Local Similarity 26.9%; Pred. No. 0.019;  
Matches 42; Conservative 7; Mismatches 42; Indels 65; Gaps 7;

Qy 14 PTVAHTSSRVAVRSVEFTCRAGLDWVCCRRLPSGRNLLEVDSLSPRVLGPR 72  
Db 112 PARASTDSDSPGR-----VRRNPA----PGSRLRLRDVPAPSP-----A 149

Qy 73 GPGLSGTGLGSM----AMRAAGGDGDSCL-----PVALGLAGAPQTPGVCRAI 117  
Db 150 GFGSSSCSPGPBRAAGTGHRSRSSGERSAOCAADRFFSGQRPAADDGPGAPARTGRGRR 209

Qy 118 WVRSSPLRLAASPTSGTYRSSAPLLEALPAPWRA 153  
Db 210 WV-----PGPWRCA 218

RESULT 5  
US-09-252-991A-18207  
; Sequence 18207; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18207

Query Match 11.4%; Score 98; DB 4; Length 162;  
Best Local Similarity 25.1%; Pred. No. 0.015;  
Matches 24; Conservative 70; Mismatches 30; Indels 30; Gaps 8;

Qy 15 TVAHRDSSRVAVRSVEFTCRAGLDWVCCRRLPSGTLPSPMAMRAAGGR 93  
Db 6 SVPRTASPAISSRSP-----ACCSATCSSACARKDWPWTSPTAWRAAKATRSLCHW 63

Qy 60 DVSLSPRLVPLVAGLAGAPQTPCVRGAIWRSIPLRAS---PTSGNTYRSSA----- 144  
Db 64 GTAFSSRASSTTSFTPTTRPACNWRA---TASCPASRSASGSPSACSGRSP 120

Qy 116 AIWRSIPLRASPTSWGTYRSSAPLLEALPGPWMAASGEWKTAT 161  
Db 121 ATAAKRATSTAAATSPCRHRSTSN-GSGLPARN-----WRAAS 160

RESULT 6  
US-09-252-991A-24790  
; Sequence 24790; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

RESULT 7  
 US -09-252-991A-22801  
 ; Sequence 22801, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO: 22801  
 ; LENGTH: 231  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-24790

Query Match 11.3%; Score 97; DB 4; Length 231;  
 Best Local Similarity 29.6%; Pred. No. 0.03; Mismatches 8; Indels 8; Gaps 7;

Matches 45; Conservative 45; Number of SEQ ID NOS: 33142

Qy 10 TNVTPVAVRTPSSVAVRSVLEFTCCRAGALDWVCA-----RRERLPSGNLEVDVLSLR 67  
 Db 8 TRIGACSSRCSSAR-----CHSACATRMPGCAARSALPSGR-----RG 47

Qy 68 VGPRAAGPGLSPGTGLGPSMAMRAAGRGSCLPVAGLQAGAPQTPGVGRAIWRSRPIURA 127  
 Db 48 SGSTAGTNWPSPTTSMSTACLARSRT---TIANIRRAPMP-----PLTV 91

Qy 120 ASPS---WGTYRSSAPPLAEALGPWNRASG 155

RESULT 8  
 US -09-252-991A-27305  
 ; Sequence 27305, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO: 27305  
 ; LENGTH: 709  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-27305

Query Match 11.2%; Score 96; DB 4; Length 709;  
 Best Local Similarity 34.5%; Pred. No. 0.16; Mismatches 6; Indels 24; Gaps 9;

Matches 50; Conservative 50; Number of SEQ ID NOS: 33142

Qy 19 RTSSRVARSLVETFCRAGALDWVCAERLPSGRNLEVDSLSR-----RLVGR 71  
 Db 96 RTPGGRPACRA --TRARRSAL--ADAGRGLPAGRQRQDALARPAPRGAERLSRSR 150

Qy 72 AGPCLSPGTGLGPVAMRAAGGRGCSLPUVALGAGAPQTPGVGRAIWRSRPIURAASPT 131  
 Db 151 AVAALRPS--RPARLARAGGAGA--PPEPGVALEPDRGAGR--WLPGG-AFPAAPEG 204

Qy 132 SWGTYRSSAPPLAEALGPWNRASG 156  
 Db 205 RLG---SRP--GAQPGFGLRACGF 223

RESULT 9  
 US -09-252-991A-17729  
 ; Sequence 17729, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO: 17729  
 ; LENGTH: 328  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-17729

Query Match 11.1%; Score 95.5; DB 4; Length 328;  
 Best Local Similarity 29.4%; Pred. No. 0.067; Mismatches 12; Indels 43; Gaps 8;

Matches 48; Conservative 48; Number of SEQ ID NOS: 33142

Qy 14 PVAHRSRSPSSRVAVRSLVETFCRAGALDWVCAERLPSGR-----NEVDVY 61  
 Db 95 PVAHRSRSPSSRVAVRSLVETFCRAGALDWVCAERLPSGR-----AR----PAGRRSGCQDHGRGPDPHSPV 133

Qy 62 SLSPRLVGPRAGPGLSPGTG-----CLPVALGLGAAQTPG 112

134 ALRPRAGADEGPKRPPAPVGGDAPALDAGGPVQRQALRRKRCQSQSLRGHPVGPGRG 193  
 Qy 113 VGRAWVRSRSLPLRASPTWGTYRSSAAPLEALPGPWNMSG 155  
 Db 194 VAEGRTRPGRPVRLRGPGGR-RSPPGIAPGFQQRTRAG 235

RESULT 10  
 US-09-252-991A-25006  
 ; Sequence 25006, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIORITY FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 25006  
 ; LENGTH: 371  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-25006

Query Match 11 1%; Score 95; DB 4; Length 371;  
 Best Local Similarity 32.6%; Pred. No. 0.080;  
 Matches 44; Conservative 10; Mismatches 53; Indels 28; Gaps 6;  
 Qy 47 RERDPSGMLNEDVDSLSPLVPGPAGPUSPGLGPMAMRAGGRGDSCLPVAGLGA 106  
 Db 31 RRDPSGMLNVRQLRQADPRAAPPAAPAGR-APG-----RQAPGRRTARPVAGTAG 82  
 Qy 107 ---AQQTGPGVGRATWVRSI----PLR---AASPTSW-----GYRSSAPLIL 146  
 Db 83 DSWAHRHKGPGDADLQSAIAHRPPGRRGPPAPSPANRAAAGTGTARTGRPRAAT 142  
 Qy 147 PGPNRMSAGFWKTAT 161  
 Db 143 PARGUSAPALRPPAT 157

RESULT 11  
 US-09-252-991A-24258  
 ; Sequence 24258, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIORITY FILING DATE: 1999-02-18  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 24258;  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-24258

Query Match 10 9%; Score 93; DB 4; Length 213;  
 Best Local Similarity 29.8%; Pred. No. 0.07;  
 Matches 54; Conservative 13; Mismatches 60; Indels 54; Gaps 10;

Qy 15 TWAHTSSSRV-AVRSILVEPTCCR----AGALDWYCARRRLPSGRNLEVTDVLSLSPR 66  
 Db 34 TPARRSSCRRLPPLALRHTOPAVRPQRPCTL----REERLRLTRRR----REERLRLTRRR 84  
 Qy 67 LVSPRAGGLSPOTLGPBMAMAG---GRDGSCLPVAGLGAQPTQPV---GRAIVNR 120  
 Db 85 LVGDQGGRAGSGDAAVVLRRQGATGRPSTRYPTAAACA-PPSAPVLRREGRGARR 143  
 Qy 121 SSIPRLA--ASPTSW-----GYRSSA-----PILLEALPGPW 150  
 Db 144 GSPGPRSPRSPGSPWPCSAPCRGRRCPCHGTAPRAGSLPRPPAFLPSPAPLPGPW 203  
 Qy 151 R 151  
 Db 204 R 204

RESULT 12  
 US-09-252-991A-23311  
 ; Sequence 23311, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIORITY FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 23311  
 ; LENGTH: 371  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-23311

Query Match 10 8%; Score 92.5%; DB 4; Length 254;  
 Best Local Similarity 29.6%; Pred. No. 0.099;  
 Matches 50; Conservative 12; Mismatches 80; Indels 27; Gaps 8;  
 Qy 13 TPTVAHTSSSRVAVRSILVEPTCC-----RAGALDWYCARRRLPS-GRNLEVDV 61  
 Db 86 TPAAPARGRADRLVRSPARLALPGSAAROARROSGATLQCARPRQAPGRRIRPGS 145  
 Qy 62 SLSPL-----RLVPGPAGPGLSPGTLSPTGSPMSMAMRAAGGRDGSCLPVAGLGAQ-P-QTPCGV 114  
 Db 146 ATAPGTAPEFRRQGRPRPAAHRWPQSPAPVAGHPGGRPEPLPGAGRHRPPVQPCVG 205  
 Qy 115 RAIWVRSIIP-----IARAASP-TSNTOTYRSSAPLALPG-PWRIASG 155  
 Db 206 -ATGLRADPGAGYGTLLRGRPEPGAAARRGTPGELRPGSPAETARG 253

RESULT 13  
 US-09-252-991A-24717  
 ; Sequence 24717, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIORITY FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 24717

LENGTH: 461  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-24717

Query Match 10.7%; Score 91.5; DB 4; Length 461;  
 Best Local Similarity 28.2%; Pred. No. 0.27;  
 Matches 42; Conservative 14; Mismatches 54; Indels 39; Gaps 7;

Qy 44 VCARRELPSGR-----NIEVDVLSSPR-----LVG-----PRASGP-----GL---- 76  
 Db 228 VAGRARRPCKLLAGAALRRRPLRRRPLRRRPLRRRPLRRPGLPAC 287  
 Qy 77 -----SPGTIGPSMAMRAAGRDGSCLPVALGLACQTPGVAIWRSIPLR 126  
 Db 288 LRVTERDRLRPPVALPEVRSGPAGSITA-PAGGAAAGKSFAGLGR-LARRSSFFS 345  
 Qy 127 AASPTSWGTYRTRSSAPLALGPWNMASG 155  
 Db 346 SATPPRATTYPAARAPMPATTEPGRSGTG 374

RESULT 14  
 US-09-252-991A-29632  
 ; Sequence 22632, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MARC J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 29632  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-29632

Query Match 10.6%; Score 91; DB 4; Length 253;  
 Best Local Similarity 32.6%; Pred. No. 0.14;  
 Matches 43; Conservative 10; Mismatches 41; Indels 38; Gaps 8;

Qy 46 ARERLPSGRNLVEY-----VSLSPRLVGPRAAGPGLSPGTLGPS-----NAMR 88  
 Db 131 ARSPGPAGRAAGGAGTGPVAGTSSAAGTSSAAGTSSAAGTSSAAGT 190

Qy 89 AAGDGDGSLP----VALGLAGAPQ-PCVGRGAIWRSIPLRAS-PTSW---GTYRS 138  
 Db 191 PAGREGSGRGPGRERRRSRGRSSAAPSGRGTGRA-----CAGGRGTSMRRRLGTGR 242

Qy 139 SAPILEALGPW 150  
 Db 243 RSP---ASFGGW 251

RESULT 15  
 US-09-252-991A-26216  
 ; Sequence 26216, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MARC J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-03-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788

Query Match 10.6%; Score 91; DB 4; Length 1027;  
 Best Local Similarity 28.9%; Pred. No. 0.86;  
 Matches 48; Conservative 8; Mismatches 56; Indels 54; Gaps 8;  
 ; SEQ ID NO 26216  
 ; LENGTH: 1027  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-26216

Query Match 10.6%; Score 91; DB 4; Length 1027;  
 Best Local Similarity 28.9%; Pred. No. 0.86;  
 Matches 48; Conservative 8; Mismatches 56; Indels 54; Gaps 8;

Qy 36 CRAGADWYCARRRL-----PSGRNLVEYDVSLSPLRVG----- 69  
 Db 179 CRGAAVP--PDRAATLAAAGAPAGGE-----SUSPRPAGGLELGRPAPAGLPSAAGRR 231  
 Qy 70 --PRAGPGLSPGTIGPSMAMRAAGGR--DGSLCIPVALGLAGAPQTGCV-----G 114  
 Db 232 ATTRAGRG--GCRGPGRAGKHGRRLPGLTARLPGRLAAAPALEGVFPWPGAAEGGLR 288  
 Qy 115 RATWVRSIPLRASPTSWGTYRTRSSAPLALGPWNMASGFWKTA 160  
 Db 289 RARWGRPVAGIRAPR--GSRKSPGFLVPCSPGTLLAAAALWPGA 331

Search completed: August 25, 2005, 03:18:11  
 Job time : 23 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

OM protein - protein search, using SW mode

Run on: August 25, 2005, 03:08:37 ; Search time 161 Seconds (without alignments)

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARILNLKKKTNVPTVAHRT.....ALPGPWRMASGFNKATMQ 164

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 10.0%  
Listing first 45 summaries:

Database : Published Applications AA:  
 1: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep:  
 2: /cgn2\_6/ptodata/1/pubpa/PTI\_NEW\_PUB.pep:  
 3: /cgn2\_6/ptodata/1/pubpa/US06\_PUB.pep:  
 4: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep:  
 5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB.pep:  
 6: /cgn2\_6/ptodata/1/pubpa/PECTUS\_PUBCOMB.pep:  
 7: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep:  
 8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB.pep:  
 9: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep:  
 10: /cgn2\_6/ptodata/1/pubpa/US09B\_PUBCOMB.pep:  
 11: /cgn2\_6/ptodata/1/pubpa/US09C\_PUBCOMB.pep:  
 12: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB.pep:  
 13: /cgn2\_6/ptodata/1/pubpa/US10A\_PUBCOMB.pep:  
 14: /cgn2\_6/ptodata/1/pubpa/US10B\_PUBCOMB.pep:  
 15: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:  
 16: /cgn2\_6/ptodata/1/pubpa/US10D\_PUBCOMB.pep:  
 17: /cgn2\_6/ptodata/1/pubpa/US10E\_PUBCOMB.pep:  
 18: /cgn2\_6/ptodata/1/pubpa/US10A\_NEW\_PUB.pep:  
 19: /cgn2\_6/ptodata/1/pubpa/US11A\_PUBCOMB.pep:  
 20: /cgn2\_6/ptodata/1/pubpa/US11A\_NEW\_PUB.pep:  
 21: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB.pep:  
 22: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB.pep:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	100.0	164	17 US-10-664-038-1	Sequence 1, Appli
2	837	97.7	160	15 US-10-150-233-1	Sequence 1, Appli
3	824	96.1	196	16 US-10-601-020-2	Sequence 2, Appli
4	824	96.1	196	16 US-10-601-020-10	Sequence 10, Appli
5	806	94.0	161	9 US-09-736-952A-1	Sequence 1, Appli
6	806	94.0	161	9 US-09-736-952A-32	Sequence 32, Appli
7	768	89.6	161	9 US-09-736-952A-3	Sequence 3, Appli
8	763	89.0	161	9 US-09-736-952A-2	Sequence 2, Appli
9	658	76.8	192	16 US-10-601-020-17	Sequence 17, Appli
10	648.5	75.7	161	9 US-09-736-952A-9	Sequence 9, Appli
11	624	72.8	161	9 US-09-736-952A-7	Sequence 7, Appli

## ALIGNMENTS

RESULT 1  
 US-10-664-038-1  
 / Sequence 1, Application US-10664038  
 / Publication No. US20050053915A1  
 / GENERAL INFORMATION:  
 / APPLICANT: VARGALIOTI, AGORITI  
 / APPLICANT: GEORGOPOLOU, UTRANIA  
 / TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH  
 / TITLE OF INVENTION: AND/OR OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE  
 / TITLE OF INVENTION: PRODUCTS  
 / FILE REFERENCE: 03495-0194-00000  
 / CURRENT APPLICATION NUMBER: US-10-664-038-1  
 / CURRENT FILING DATE: 2003-07-17  
 / PRIOR APPLICATION NUMBER: US-09-644-987  
 / PRIOR FILING DATE: 2000-08-24  
 / PRIOR APPLICATION NUMBER: 60/151,074  
 / PRIOR FILING DATE: 1999-08-27  
 / NUMBER OF SEQ ID NOS: 16  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO: 1  
 / LENGTH: 164  
 / TYPE: PRT  
 / ORGANISM: Hepatitis C virus  
 / US-10-664-038-1

Query Match Score 857; DB 17; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-70;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARILNLKKKTNVPTVAHRTSSRVAVRSLVFTCCRAGALDWVCCRERLPSGRNLLEVD 60  
 Db 1 ARILNLKKKTNVPTVAHRTSSRVAVRSLVFTCCRAGALDWVCCRERLPSGRNLLEVD 60

Qy 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQPGVGRAIWR 120 ; TYPE: PRT  
 Db 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQPGVGRAIWR 120 ; ORGANISM: Hepatitis C virus  
 US-10-601-020-2

Qy 121 SSIPRAASPTSWGTYRSAPLLEALPQPMASGFWKATMQ 164 ; Query Match 96.1%; Score 824; DB 16; Length 196;  
 Db 121 SSIPRAASPTSWGTYRSAPLLEALPQPMASGFWKATMQ 164 ; Best Local Similarity 96.3%; Pred. No. 2.2e-67;  
 Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 2

US-10-150-283-1  
 Sequence 1, Application US/10150283  
 Publication No. US20030219407A1  
 GENERAL INFORMATION:  
 APPLICANT: Ding, Shou-wei  
 APPLICANT: Li, Hong-wei  
 APPLICANT: Li, Wan-xiang  
 TITLE OF INVENTION: RNA Silencing in Animals as an Antiviral Defense  
 PILE REFERENCE: 03070-12410US  
 CURRENT APPLICATION NUMBER: US/10/150,283  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 160  
 TYPE: PRT  
 ORGANISM: Hepatitis C virus  
 FEATURE:  
 OTHER INFORMATION: P protein

US-10-150-283-1

Query Match 97.7%; Score 837; DB 15; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-68; ; Indels 0; Gaps 0;  
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARILNLKKTNVTPYAHRTSSRVAVRSLVEFTCCRAGLDWVCAERRLPSGRNLEVD 60 ; Query Match 96.1%; Score 824; DB 16; Length 196;  
 Db 1 ARILNLKKTNVTPYAHRTSSRVAVRSLVEFTCCRAGLDWVCAERRLPSGRNLEVD 60 ; Best Local Similarity 96.3%; Pred. No. 2.2e-67;  
 Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 4

US-10-601-020-10  
 Sequence 10, Application US/10601020  
 Publication No. US20040156862A1  
 GENERAL INFORMATION:  
 APPLICANT: Branch, Andrea D.  
 APPLICANT: Walewski, Jose L.  
 APPLICANT: Stump, Dechard D.  
 TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 FILE REFERENCE: RII-003CPUSCN  
 CURRENT APPLICATION NUMBER: US/10/601,020  
 CURRENT FILING DATE: 2003-06-20  
 PRIORITY APPLICATION NUMBER: US 09/719277  
 PRIORITY FILING DATE: 2001-04-13  
 PRIORITY APPLICATION NUMBER: US 60/088670  
 PRIORITY FILING DATE: 1998-06-09  
 PRIORITY APPLICATION NUMBER: US 60/089138  
 PRIORITY FILING DATE: 1998-06-11  
 PRIORITY APPLICATION NUMBER: PCT/US99/12929  
 PRIORITY FILING DATE: 1999-06-09  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 10  
 LENGTH: 196  
 TYPE: PRT  
 ORGANISM: Hepatitis C virus

US-10-601-020-10

Query Match 96.1%; Score 824; DB 16; Length 196;  
 Best Local Similarity 96.3%; Pred. No. 2.2e-67;  
 Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQPGVGRAIWR 120 ; Query Match 96.1%; Score 824; DB 16; Length 196;  
 Db 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQPGVGRAIWR 120 ; Best Local Similarity 96.3%; Pred. No. 2.2e-67;  
 Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 3

US-10-601-020-2  
 Sequence 2, Application US/10601020  
 Publication No. US20040156862A1  
 GENERAL INFORMATION:  
 APPLICANT: Branch, Andrea D.  
 APPLICANT: Walewski, Jose L.  
 APPLICANT: Stump, Dechard D.  
 TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 FILE REFERENCE: RII-003CPUSCN  
 CURRENT APPLICATION NUMBER: US/10/601,020  
 CURRENT FILING DATE: 2003-06-20  
 PRIORITY APPLICATION NUMBER: US 09/719277  
 PRIORITY FILING DATE: 2001-04-13  
 PRIORITY APPLICATION NUMBER: US 60/088670  
 PRIORITY FILING DATE: 1998-06-09  
 PRIORITY APPLICATION NUMBER: US 60/089138  
 PRIORITY FILING DATE: 1998-06-11  
 PRIORITY APPLICATION NUMBER: PCT/US99/12929  
 PRIORITY FILING DATE: 1999-06-09  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 196

RESULT 5

US-09-736-959A-1  
 Sequence 1, Application US/09736959A  
 Publication No. US2002007641A1  
 GENERAL INFORMATION:  
 APPLICANT: University of Southern California  
 APPLICANT: OU, Jing-Hsuing

```

RESULT 7
US-09-736-959A-3
; Sequence 3, Application US/09736959A
; Patent No. US20070641A1
; GENERAL INFORMATION:
;   APPLICANT: University of Southern California
;   APPLICANT: OU, Jing-Hsiung
;   APPLICANT: XU, Zhenming
;   TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
;   FILE REFERENCE: 13761-742
;   CURRENT APPLICATION NUMBER: US/09/736,959A
;   CURRENT FILING DATE: 2000-12-14
;   PRIOR APPLICATION NUMBER: US 60/170,835
;   PRIOR FILING DATE: 1999-12-14
;   NUMBER OF SEQ ID NOS: 32
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
; FEATURE:
; OTHER INFORMATION: Exemplary P17 sequences

US-09-736-959A-3
Query Match 89.6%; Score 768; DB 9; Le
Best Local Similarity 95.5%; Pred. No. 2.3e-62;
Matches 147; Conservative 2; Mismatches 5;
Qy 7 KCKCTNVTPTVYHRTSSSRVAVRSLEPTTCRAGALDWYCAR
Db 8 QKRPNTVPTVYHRTSSSRVAVRSLEPTTCRAGALDWYCAR
Qy 67 LVGPRGPGLGPGLGPSSMAMRAAGRDGSCLPVAGLGLAGA
Db 68 LVDPRGPGLGPGLGPSSMAMRAAGRDGSCLPVAGLGLAGA
Qy 127 AASPTSGTSSAPLLEALGPWMAASGFWKTA 160
Db 128 AASPTSGTSSAPLLEALGPWMAASGFWKTA 161

RESULT 8
US-09-736-959A-2
; Sequence 2, Application US/09736959A
; Patent No. US20070641A1
; GENERAL INFORMATION:
;   APPLICANT: University of Southern California
;   APPLICANT: OU, Jing-Hsiung
;   APPLICANT: XU, Zhenming
;   TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
;   FILE REFERENCE: 13761-742
;   CURRENT APPLICATION NUMBER: US/09/736,959A
;   CURRENT FILING DATE: 2000-12-14
;   PRIOR APPLICATION NUMBER: US 60/170,835
;   PRIOR FILING DATE: 1999-12-14
;   NUMBER OF SEQ ID NOS: 32
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
; FEATURE:
; OTHER INFORMATION: Exemplary P17 sequences

US-09-736-959A-2
Query Match 89.0%; Score 763; DB 9; Le
Best Local Similarity 94.8%; Pred. No. 6.7e-62;
Matches 146; Conservative 2; Mismatches 6;
Qy 7 KKKCTNVTPTVYHRTSSSRVAVRSLEPTTCRAGALDWYCAR
;
```

RESULT 9

Db 8 ORKPNVTPPTVAHRTSSSWAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEVDVLSLSPR 67  
 Qy 67 LVGPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIYRSSIPL 126  
 Db 68 HVGPPAGPGLSPGTLGPSPMAMRVAGWDSCCLPVALGLAGAPOTPGVGRAIYRSSIPL 127

Qy 127 AASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 160  
 Db 128 AASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 161

Db US-10-601-020-9

Query Match 75.7%; Score 648.5; DB 16; Length 195;  
 Best Local Similarity 78.7%; Pred. No. 2.5e-51;  
 Matches 129; Conservative 6; Mismatches 28; Indels 1; Gaps 1;

Qy 1 ARILNKKTNTPVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEVD 60  
 Db 1 AQILNKKENPTVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEVG 60

Qy 61 VSLSPRLVPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIYWR 120  
 Db 61 ASLSPRLVPRAGPGLSPGTLGPSPMAMRVAGQDGSCPPAALGLGAPMTPGGPAIYWR 120

Qy 121 SSIPRLRASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 164  
 Db 121 SSIPRLRASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 163

RESULT 11

US-09-36-959A-7

Qy Sequence 7, Application US/09736959A  
 Db Patent No. US20020076415A1

GENERAL INFORMATION:  
 APPLICANT: Branch, Andrea D.  
 APPLICANT: Stump, DeCherd D.  
 TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 FILE REFERENCE: RII-003CPUSCN

CURRENT APPLICATION NUMBER: US/10/601,020  
 CURRENT FILING DATE: 2003-06-20  
 PRIOR APPLICATION NUMBER: US 09/719277  
 PRIOR FILING DATE: 2001-04-13  
 PRIOR APPLICATION NUMBER: US 60/088670  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: US 60/089138  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: PCT/US99/12929  
 NUMBER OF SEQ ID NOS: 17  
 SEQ ID NO 17  
 LENGTH: 192  
 TYPE: PRT  
 ORGANISM: Hepatitis C virus  
 US-10-601-020-17

Query Match 76.8%; Score 658; DB 16; Length 192;  
 Best Local Similarity 77.4%; Pred. No. 3.3e-52;  
 Matches 127; Conservative 9; Mismatches 26; Indels 2; Gaps 1;

Qy 1 ARILNKKTNTPVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEVD 60  
 Db 1 AQILNKKENPTVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEG 60

Qy 61 VSLSPRLVPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIYWR 120  
 Db 61 DNLSPRAGPGLSPGTLGPSPMAMRVGGQDGSCPPAALGLGAPMTPGGPAIYWR 120

Qy 121 SSIPRLRASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 164  
 Db 121 SSIPRLRASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 162

RESULT 10

US-10-601-020-9

Qy Sequence 9, Application US/10601020  
 Db Publication No. US20040156862A1

GENERAL INFORMATION:  
 APPLICANT: Branch, Andrea D.  
 APPLICANT: Stump, DeCherd D.  
 APPLICANT: Walewski, Jose L.

TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 FILE REFERENCE: RII-003CPUSCN

CURRENT APPLICATION NUMBER: US/10/601,020  
 CURRENT FILING DATE: 2003-06-20  
 PRIOR APPLICATION NUMBER: US 09/719277  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: US 60/088670  
 PRIOR FILING DATE: 2001-04-13  
 PRIOR APPLICATION NUMBER: US 60/089138

Query Match 72.8%; Score 624; DB 9; Length 161;  
 Best Local Similarity 77.3%; Pred. No. 3.5e-19;  
 Matches 119; Conservative 7; Mismatches 28; Indels 0; Gaps 0;

Qy 7 KCKTNVTPVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEVDVLSLSPR 66  
 Db 8 QRKPNVTPVARTSSRVAHRTSSRVAVRSLVFTCCRAGALDWVCAAGRLPSGRNLVEGDNLSPR 67

Qy 67 LVGPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIYRSSIPLR 126  
 Db 68 LADPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVALGLAGAPOTPGVGRAIYRSSIPLS 127

Qy 127 AASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 160  
 Db 128 AASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 161

RESULT 12

US-10-601-020-13

Qy Sequence 13, Application US/10601020  
 Db Publication No. US20040156862A1

GENERAL INFORMATION  
 / APPLICANT: Branch, Andrea D.  
 / APPLICANT: Walewski, Jose L.  
 / APPLICANT: Stump, Decherd D.  
 / TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 / FILE REFERENCE: RAI-003CPUSCN  
 / CURRENT APPLICATION NUMBER: US/10/601,020  
 / CURRENT FILING DATE: 2003-06-20  
 / PRIOR APPLICATION NUMBER: US 09/719277  
 / PRIOR FILING DATE: 1998-04-13  
 / PRIOR APPLICATION NUMBER: US 60/088670  
 / PRIOR FILING DATE: 1998-06-09  
 / PRIOR APPLICATION NUMBER: US 60/089138  
 / PRIOR FILING DATE: 1998-06-11  
 / PRIOR APPLICATION NUMBER: PCT/US99/12929  
 / PRIOR FILING DATE: 1999-06-09  
 / NUMBER OF SEQ ID NOS: 17  
 / SOFTWARE: PatentIn Ver. 2.0  
 / SEQ ID NO: 13  
 / LENGTH: 195  
 / TYPE: PRT  
 / ORGANISM: Hepatitis C virus  
 / US-10-601-020-13

Query Match 68.1%; Score 583.5; DB 16; Length 195;  
 Best Local Similarity 73.8%; Pred. No. 2.2e-45;  
 Matches 121; Conservative 2; Mismatches 40; Indels 1; Gaps 1;

Db 1 ARILNLKPKTKNTVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 60  
 1 ARILNLKPKPTVAAHRTSSRVAVRLAFTCCRAGAPEWCARRGRPSGRNLVEVD 60

Qy 61 VSLSPRLVGRPAGPGLSPGTLSLPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 120  
 61 VSLSPRPADPREGPGRSPGTLSLPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 120

Db 121 SSIPRLAASPSPSWGTYRSAPLLEALPGLPWRMASGPKATMQ 164  
 121 SSIPRLAASPSPSWGTYRSAPLLEALPGLPWRMASGPKATMQ 164

Qy 121 SSIPRLAASPSPSWGTYRSAPLLEALPGLPWRMASGPKATMQ 164  
 121 SSIPRLAASPSPSWGTYRSAPLLEALPGLPWRMASGPKATMQ 164

Db 121 SSIPRLAASPSPSWGTYRSAPLLEALPGLPWRMASGPKATMQ 164

RESULT 13  
 US-09-736-959A-5  
 / Sequence 5, Application US/09736959A  
 / Patent No. US20020076415A1  
 / GENERAL INFORMATION:  
 / APPLICANT: University of Southern California  
 / APPLICANT: OU, Jing-Hsiung  
 / APPLICANT: XU, Zheming  
 / FILE REFERENCE: 13-61-742  
 / TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
 / CURRENT APPLICATION NUMBER: US/09/736,959A  
 / CURRENT FILING DATE: 2000-12-14  
 / PRIOR APPLICATION NUMBER: US 60/170,835  
 / PRIOR FILING DATE: 1999-12-14  
 / NUMBER OF SEQ ID NOS: 32  
 / SEQ ID NO: 5  
 / LENGTH: 143  
 / TYPE: PRT  
 / ORGANISM: Hepatitis C Virus  
 / OTHER INFORMATION: Exemplary P17 sequences  
 / US-09-736-959A-5

Query Match 64.4%; Score 552; DB 9; Length 143;  
 Best Local Similarity 79.3%; Pred. No. 1.2e-42;  
 Matches 107; Conservative 5; Mismatches 23; Indels 0; Gaps 0;

Db 127 AASPTSWGTYRSSAP 141  
 128 AASPTSWGTYRSSAP 142

RESULT 14  
 US-09-736-959A-9  
 / Sequence 9, Application US/09736959A  
 / Patent No. US20020076415A1  
 / GENERAL INFORMATION:  
 / APPLICANT: University of Southern California  
 / APPLICANT: OU, Jing-Hsiung  
 / APPLICANT: XU, Zheming  
 / FILE REFERENCE: 13-61-742  
 / TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
 / CURRENT APPLICATION NUMBER: US/09/736,959A  
 / CURRENT FILING DATE: 2000-12-14  
 / PRIOR APPLICATION NUMBER: US 60/170,835  
 / PRIOR FILING DATE: 1999-12-14  
 / NUMBER OF SEQ ID NOS: 32  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO: 9  
 / LENGTH: 143  
 / TYPE: PRT  
 / ORGANISM: Hepatitis C Virus  
 / FEATURE:  
 / OTHER INFORMATION: Exemplary P17 sequences  
 / US-09-736-959A-9

Query Match 63.9%; Score 548; DB 9; Length 143;  
 Best Local Similarity 78.5%; Pred. No. 2.8e-42;  
 Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

Db 7 KKKTNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 66  
 8 QRKPNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 67

Qy 67 LVGRPAGPGLSPGTLSLPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 126  
 68 LADPRAAGPGLSPGTLSLPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 127

Db 127 AASPTSWGTYRSSAP 141  
 128 AASPTSWGTYRSSAP 142

RESULT 15  
 US-09-736-959A-8  
 / Sequence 8, Application US/09736959A  
 / Patent No. US20020076415A1  
 / GENERAL INFORMATION:  
 / APPLICANT: University of Southern California  
 / APPLICANT: OU, Jing-Hsiung  
 / APPLICANT: XU, Zheming  
 / FILE REFERENCE: 13-761-742  
 / TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
 / CURRENT APPLICATION NUMBER: US/09/736,959A  
 / CURRENT FILING DATE: 2000-12-14  
 / PRIOR APPLICATION NUMBER: US 60/170,835  
 / PRIOR FILING DATE: 1999-12-14  
 / NUMBER OF SEQ ID NOS: 32  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO: 8  
 / LENGTH: 143  
 / TYPE: PRT  
 / ORGANISM: Hepatitis C Virus  
 / FEATURE:  
 / OTHER INFORMATION: Exemplary P17 sequences  
 / US-09-736-959A-8

Query Match 7 KKKTNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 66  
 8 QRKPNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 67

Qy 7 KKKTNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 66  
 8 QRKPNVPTVAAHRTSSRVAVRLPGLGAGPDTGGSCLPVAGLGLAATGQ 67

Query Match 63.8%; Score 547; DB 9; Length 143;  
Best Local Similarity 78.5%; Pred. No. 3; 5e-42;  
Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

---

Qy 7 KKCTINVTPTVAHRTSSSRVAVRSILVEFTCRAGALDWYCARRERLPSGRNLLEVDSLSPR 66  
Db 8 QRKSNVTPTAAHRTSSRAVRSILVEFTCRAGAPGWCARLGRPSGRNLVEGDNLSPR 67

Qy 67 LVGRAGPGLSPGTLGPSNAMRAAGGRDSCULPVALGLAGAPQTPGVRCAINRSSIPLR 126  
Db 68 LASPRAGPGLSPGTLGPSNAMRAWGQDGSCHPAAAPGLVGAPKTPGVCRVIVRSSIPSH 127

Qy 127 AASPTSWGYRSSAP 141.  
Db 128 AASPTSWGYRSSAP 142

Search completed: August 25, 2005, 03:29:13  
Job time : 162 secs